

Enhancements to the USGS Landsat Level 2 surface temperature and emissivity product for Collection 3 reprocessing[☆]

Saeed Arab^{a,*}, Chase Mueller^a, Christopher J. Crawford^b, Danika Wellington^a, Ray Dittmeier^a, Glynn Hulley^c, Aaron Gerace^d, Calli Jenkerson^a, Simon Hook^c, Amirhossein Hassanzadeh^d, Rehman Eon^d, Cody Anderson^b, Pat Scaramuzza^a, Frank M. Göttsche^e

^a KBR, Contractor to the U.S. Geological Survey (USGS), Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, 57198, USA

^b U.S. Geological Survey (USGS), Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, 57198, USA

^c Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109, USA

^d Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, 14623, USA

^e IMKASF, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

ARTICLE INFO

Edited by Zhe Zhu

Keywords:

Landsat
Surface temperature and emissivity
Thermal infrared
Analysis ready data
Collection 3

ABSTRACT

The Landsat program provides the longest continuous global record of thermal infrared observations of the Earth's surface, underpinning critical applications in climate monitoring, water resources, ecosystem dynamics, urban heat analysis, and natural hazard assessment. The release of a global inventory of Landsat Collection 2 Level 2 surface temperature products by the U.S. Geological Survey (USGS) marked a major milestone in operational provision of Landsat thermal infrared analysis-ready data. Ongoing validations and community uses of Collection 2 have identified opportunities to further improve accuracy, uncertainty characterization, and emissivity correction across diverse atmospheric and surface conditions. In preparation for the planned Landsat Collection 3 reprocessing of the Landsat data record in the late 2020s, the USGS is implementing a coordinated set of enhancements to the Level 2 surface temperature products. These include revised emissivity estimation that leverages external datasets, improved atmospheric characterization and uncertainty propagation, expanded dynamic range for high temperature targets, consideration of split window atmospheric correction algorithm for Landsat 8 and 9, and decoupling of thermal infrared processing from visible to shortwave infrared constraints to enable surface temperature retrievals under low or no solar illumination conditions. These changes are designed to improve product quality and consistency across the Landsat record. Beyond near-term performance gains, the Collection 3 design establishes a scalable processing architecture to accommodate the expanded spectral and radiometric measurement capabilities of the forthcoming Landsat 10 mission. By preserving continuity across the Landsat 4–9 record while enabling future algorithm evolution, Landsat Collection 3 will provide a foundation for long-term, multi-decadal Earth system thermal infrared observations.

1. Introduction

Landsat satellites have provided continuous measurements of thermal infrared (TIR) emission from the Earth's surface for more than four decades, beginning in 1982 (Wulder et al., 2022). The foundational products derived from TIR measurements are the atmospherically corrected Level-2 surface temperature and spectral emissivity. Atmospheric correction of remotely sensed TIR data involves compensating for atmospheric effects on apparent radiance to solve for surface emitted

thermal radiance and subsequently derive surface temperature (Jimenez-Munoz et al., 2014; Malakar et al., 2018). This process typically requires accurate characterization of atmospheric state, together with knowledge or retrieval of surface emissivity (Barsi et al., 2003).

In 2018, the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, in collaboration with the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) and Rochester Institute of Technology (RIT), provisionally released Landsat Collection 1 (C1) Level 2 surface temperature for

[☆] This article is part of a Special issue entitled: 'Medium and High Resolution Satellite' published in Remote Sensing of Environment.

* Corresponding author.

E-mail address: sarab@contractor.usgs.gov (S. Arab).

<https://doi.org/10.1016/j.rse.2026.115563>

Received 1 April 2026; Received in revised form 26 June 2026; Accepted 10 July 2026

Available online 31 July 2026

0034-4257/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Landsat 4–8 acquisitions over U.S. regions, distributed as tiled Landsat U.S. Analysis Ready Data (ARD) products (Dwyer et al., 2018; Malakar et al., 2018). With the release of Landsat Collection 2 (C2) in 2020, the USGS provided the first global inventory of atmospherically corrected Level 2 surface temperature products for Landsat 4–9 (Earth Resources Observation and Science (EROS) Center, 2020a, 2020b, 2020c; Crawford et al., 2023). While Landsat TIR observations have spatial resolutions ranging from 60 to 120 m and differing spectral response functions across missions, C2 provides a consistent and continuous surface temperature record from 1982 to present (2026), at a common 30-m grid across Landsat 4 through 9. Atmospheric effects in C2 surface temperature are corrected using a single channel algorithm applied to TIR bands centered at approximately 11.4 μm (Landsat 4–7) and 10.8 μm (Landsat 8–9), while emissivity effects are corrected for using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset (GED) (Cook et al., 2014; Malakar et al., 2018; Borbas et al., 2018; Hulley et al., 2015).

The global availability of the C2 surface temperature product enabled comprehensive validation across diverse atmospheric and surface conditions, allowing overall algorithm performance and product maturity to be assessed, and revealing key strengths and limitations. Assessments by the USGS, the surface temperature user community, and other stakeholders indicated the need for algorithmic enhancements and processing improvements in the next major Landsat data record reprocessing. The Collection 3 (C3) reprocessing of the Landsat archive, currently planned for the late 2020s, will include multiple enhancements for retrieving surface temperature and emissivity from Landsat 4–9 observations while also establishing a pathway for integrating Landsat 10 observations into the data record.

In this paper, the characteristics of the Landsat C2 Level 2 surface temperature product are reviewed, and the planned Level 2 surface temperature algorithmic and processing updates for C3 are described in detail. We also demonstrate the improvements in accuracy and temporal stability, address TIR remote sensing community feedback on the C2 Level 2 surface temperature product, and briefly outline the algorithmic and processing approaches envisioned for Landsat 10 TIR observations in the 2030s.

2. Background

The Landsat program's first attempt at TIR imaging was carried out by the Landsat 3 Multispectral Scanner System (MSS) in 1978; however, its thermal instrument failed shortly after the launch (Southworth, 1983; Loveland and Dwyer, 2012). Operational TIR observations for the Landsat program therefore began with the Thematic Mapper (TM) instruments onboard Landsat 4 and 5, which acquired thermal observations in a single spectral band spanning 10.4–12.5 μm with a ground sampling distance (GSD) of 120 m (Arvidson et al., 2013). The Landsat 7 Enhanced Thematic Mapper Plus (ETM+) retained this instrument design but doubled the number of TIR detectors (eight instead of four), improving the GSD to 60 m (Schott et al., 2012). Dual-channel TIR imaging was introduced with the Landsat 8 Thermal Infrared Sensor (TIRS) and continued with a nearly identical instrument onboard Landsat 9. The Landsat 8 and 9 TIRS instruments acquire data in two spectral bands centered at 10.8 and 12.0 μm with a 100-m GSD (Barsi et al., 2014). This spatial resolution was coarser than the ETM+ TIR band to meet the more stringent noise equivalent change in temperature (NEdT) requirements associated with the pushbroom scan architecture of Landsat 8 and 9 TIRS instruments (Irons et al., 2012; Loveland and Irons, 2016).

Landsat 10, under development by NASA and USGS at the time of writing, is intended to substantially advance the spectral and spatial capabilities of Landsat TIR observations while preserving the program's core objective of maintaining a well-calibrated and radiometrically consistent record (Roy et al., 2026). Landsat 10 mission design includes five TIR spectral bands, adding three new channels centered at 8.3, 8.6,

and 9.1 μm to emulate the long-wavelength spectral coverage of ASTER (Yamaguchi et al., 1998) and the ECOSystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) (Hulley et al., 2017). These additional bands would enable spectral surface emissivity retrievals along with improved accuracy in surface temperature retrievals (Gillespie et al., 1998; Hulley and Hook, 2010; Hulley et al., 2021) and aid in discriminating silicate minerals, characterizing soil and rock types, and supporting geological mapping, resource exploration, and environmental hazard assessments (Hook et al., 1992; Ramsey and Christensen, 1998; Wu et al., 2019; Rabuffi et al., 2025). The two spectral bands centered at 11.3 and 12.0 μm will provide continuity with Landsat 8 and 9 TIRS observations. Radiometrically stable thermal instruments ensure continuous reliable measurements, providing robust long-term consistency for surface temperature retrievals across Landsat missions.

Landsat TIR instruments are calibrated pre- and post-launch to ensure radiometric, spatial, and spectral performance. Prior to launch, the Landsat 8 and 9 TIRS instruments were characterized in a thermal vacuum chamber to simulate the expected on-orbit operating conditions (Thome et al., 2011; Wenny et al., 2015; Pearlman et al., 2018). After launch, a combination of internal and vicarious calibration techniques is routinely applied to monitor sensor performance, correct for instrument degradation over the life of missions, and maintain a consistent and continuous data record (Barsi et al., 2014; Barsi et al., 2017; Barsi et al., 2022). Advances in instrument design and calibration techniques have enabled Landsat TIR measurements to meet or exceed design requirements, particularly with respect to radiometric noise and stability (Montanaro et al., 2014a; Montanaro et al., 2022). Design modifications implemented to mitigate the stray light issue identified in Landsat 8 TIRS have resulted in substantially improved radiometric performance for Landsat 9 (Montanaro et al., 2022). On-orbit calibration using internal blackbody and deep-space views indicates noise levels below 0.1 K, with radiometric stability of 0.1% within a power cycle (Barsi et al., 2022; Montanaro et al., 2024). In addition, quarterly instrument stability and relative gain reports issued by the USGS indicate that NEdT performance for Landsat 8 and 9 TIRS is consistently well within the 0.4 K mission requirements (e.g., Haque et al., 2025).

Spaceborne TIR observations support a broad range of Earth science applications, including evapotranspiration mapping (Anderson et al., 2012; Senay et al., 2013; Senay et al., 2023), urban heat island (UHI) monitoring (Zhou et al., 2018; Shreevastava et al., 2025; Hulley et al., 2026), surface material composition and mineral mapping (Vaughan et al., 2005; Rabuffi et al., 2025), geological mapping (Hook et al., 1992; Cooper et al., 2002), wildfire detection (Vlassova et al., 2014; Allison et al., 2016), volcanological assessments (Vaughan and Hook, 2006), analyses of lake thermal dynamics (Schott, 1986; Livo et al., 2007), regional water resource evaluation (Thenkabail et al., 2010), and investigations of local climate trends (Cawse-Nicholson et al., 2026). With a consistent sun-synchronous orbit and nominal mean local overpass time of 10:00 to 10:30 am, the Landsat program provides a unique long-term TIR data record that is particularly well suited for climatological trend analyses and monitoring of surface energy balance processes at local to regional scales (Guillevic et al., 2019). These applications typically rely on atmospherically corrected Level 2 surface temperature and spectral emissivity products.

The 2020 release of USGS Landsat C2 surface temperature on a near global scale joined only a few satellite-derived temperature datasets of comparable resolution. Prior to Landsat C2, the ASTER Level 2 Surface Kinetic Temperature product (AST_08) provided global surface temperature at 90 m every 16 days, and ECOSTRESS provided surface temperature every 3–5 days at 70 m. Other global imager missions such as the Moderate Resolution Imaging Spectroradiometer (MODIS), the Visible Infrared Imaging Radiometer Suite (VIIRS), and the Sentinel-3 Sea and Land Surface Temperature Radiometer (SLSTR) produce surface temperature products at coarse resolutions, but on daily timesteps. Looking ahead, several planned missions in the late 2020s, such as the

Thermal InfraRed Imaging Satellite for High-resolution Natural resource Assessment (TRISHNA) instrument from France's Centre National d'Etudes Spatiales (CNES) and the Indian Space Research Organization (ISRO) (Lagouarde et al., 2018), the Copernicus Land Surface Temperature Monitoring (LSTM) mission (Bernard et al., 2023), as well as missions planned for the 2030s such as Explorer for Artemis Geology Lunar and Earth (EAGLE)-TIR and Landsat 10, are expected to further expand the availability of medium resolution TIR observations and surface temperature and emissivity products (Cawse-Nicholson et al., 2026).

The anticipated reprocessing of the Landsat data record represents an opportunity to reformulate the Level 2 atmospherically corrected TIR data products to better align with current and future missions described above. As such, USGS EROS is considering renaming the Level 2 atmospherically corrected TIR products to "Surface Temperature and Emissivity", abbreviated as STE. This naming convention aligns Landsat terminology with comparable products from other medium- to moderate-resolution TIR missions. For example, MODIS provides a daily land surface temperature (LST) and emissivity product (MOD11; Wan and Dozier, 1996) and a newer physics-based LST&E product (MOD21; Hulley and Hook, 2010). The ECOSTRESS instrument aboard the International Space Station (ISS) similarly provides an atmospherically corrected LST&E product (ECO2LSTE; Hulley et al., 2021). The planned EAGLE-TIR mission was also initially designed to deliver a Level 2 LST&E product (Cawse-Nicholson et al., 2021). Although current Landsat missions do not directly retrieve surface spectral emissivity, since at least 3 bands are required in the TIR (at least one between 8 and 9 μm , and at least two between 10 and 12 μm), spectral and temporal adjustment techniques applied to external emissivity sources provides required surface emissivity estimates at the time of Landsat overpass. As noted, the potential addition of three TIR spectral bands between 8 and 9.5 μm on the Landsat 10 mission would enable direct surface emissivity retrievals for the first time in Landsat mission history. Adoption of the STE nomenclature for Collection 3 will provide consistent Level 2 TIR product terminology across Landsat 4–9 missions, as well as Landsat 10.

Most importantly, preparation for Collection 3 reprocessing provides the opportunity to reassess the algorithmic approaches to translate Level 1 at-sensor thermal radiance into atmospherically corrected Level 2 surface temperature. The following section describes the currently operational Landsat Collection 2 surface temperature product to set the stage for improvements and analysis in support of Landsat Collection 3 discussed in the remainder of the paper.

3. Landsat Collection 2 surface temperature

The initial release of the C2 Landsat Level 2 surface temperature product in 2020 included Landsat 4–8 acquired data with Landsat 9 being added shortly after the launch in September 2021. Newly acquired Landsat 8 and 9 image data continue to be processed and added to C2 in the forward stream as required atmospheric auxiliary inputs become available for Level 2 processing. In addition to the primary surface temperature raster layer, which reports surface skin temperature in units of Kelvin, the C2 product provides per-pixel estimates of surface temperature uncertainty, spectral emissivity and its associated standard deviation, surface thermal radiance, and multiple atmospheric characterization variables that include atmospheric transmittance, upwelling radiance, and downwelling radiance, as well as the distance to the nearest cloud detected pixel. Table 1 summarizes the per-pixel information included in the Landsat C2 surface temperature product, while Fig. 1 presents an example of Landsat 9 surface temperature alongside the corresponding true color Level 2 surface reflectance.

A comprehensive overview of the C2 atmospheric correction algorithm, emissivity retrieval, atmospheric auxiliary data, uncertainty estimation, and validation is provided in Section 3.1 through Section 3.5. Shortly after the release of C2, and informed by user community feedback, USGS EROS initiated research and development activities to

Table 1

Summary of per-pixel information included in the Landsat Collection 2 (C2) surface temperature product.

Band name	Units	Description/purpose
Surface Temperature	K	Primary geophysical quantity derived from Landsat 4–5 TM and Landsat 7 ETM+ Band 6 and Landsat 8–9 TIRS Band 10
Surface Temperature Uncertainty	K	Per-pixel measure of surface temperature uncertainty derived from a standard error propagation model
Emissivity	Unitless	Spectral emissivity derived from ASTER GED and spectrally and temporally adjusted to the date of the Landsat overpass
Emissivity Standard Deviation	Unitless	Intermediate layer used in uncertainty propagation model
Thermal Radiance	W/($\text{m}^2\cdot\text{sr}\cdot\mu\text{m}$)	Level 1 thermal infrared radiance used as input to the single channel equation
Upwelling Radiance	W/($\text{m}^2\cdot\text{sr}\cdot\mu\text{m}$)	Atmospheric characterization information derived from radiative transfer model
Downwelling Radiance	W/($\text{m}^2\cdot\text{sr}\cdot\mu\text{m}$)	Atmospheric characterization information derived from radiative transfer model
Atmospheric Transmittance	Unitless	Atmospheric characterization information derived from radiative transfer model
Pixel Distance to Cloud	km	Intermediate layer used in uncertainty propagation model

TM = Thematic Mapper, ETM+ = Enhanced Thematic Mapper Plus, TIRS = Thermal Infrared Sensor, ASTER = Advanced Spaceborne Thermal Emission and Reflection Radiometer, GED = Global Emissivity Dataset.

improve the Level 2 surface temperature product. These efforts focused on addressing known limitations and issues identified in the global C2 product, as described in Section 3.6, and expanded to improve the overall product accuracy, precision, and quality. The primary algorithmic and processing enhancements planned for C3 processing of Level 2 surface temperature and spectral surface emissivity are described in Section 4.

3.1. Atmospheric correction algorithm

During the development phase of Landsat C2, two atmospheric correction algorithms were evaluated for processing Landsat TIR image data: a single channel algorithm adapted from Cook et al. (2014), and a split window algorithm from Gerace et al. (2020). For the historical Landsat 4–5 TM and Landsat 7 ETM+ data records, the single channel algorithm was the only viable option because only one TIR band was available. In contrast, for the dual-channel Landsat 8–9 TIRS observations, either algorithm could be applied. Following extensive analysis and characterization of the differences between these techniques, USGS selected the single channel algorithm for processing the Landsat 4–9 TIR data record for three primary reasons. First, applying a consistent atmospheric correction algorithm to Landsat 4–9 TIR observations minimized the risk of algorithm-induced discontinuities over the multidecadal record. This was an especially important consideration because C2 represented the first global-generated surface temperature product from the Landsat data record. Second, the single channel algorithm had a longer development and validation history because it was provisionally implemented in Landsat C1 U.S. ARD over the United States (Malakar et al., 2018; Dwyer et al., 2018). Concomitant with this greater processing maturity was the availability of per-pixel surface temperature uncertainty estimates within the single channel algorithm implementation, providing additional quality assurance (Laraby and Schott, 2018). Third, applying the single channel algorithm to Landsat 8 TIRS Band 10 avoided aliasing artifacts associated with physical layout of detectors and residual stray light effects affecting Band 11. The physical offset between the Landsat 8 TIRS Bands 10 and 11 on the focal plane, combined with upsampling the native 100-m observations to 30 m, produced along-track striping and aliasing artifacts in the difference and average apparent temperature fields used by the split window

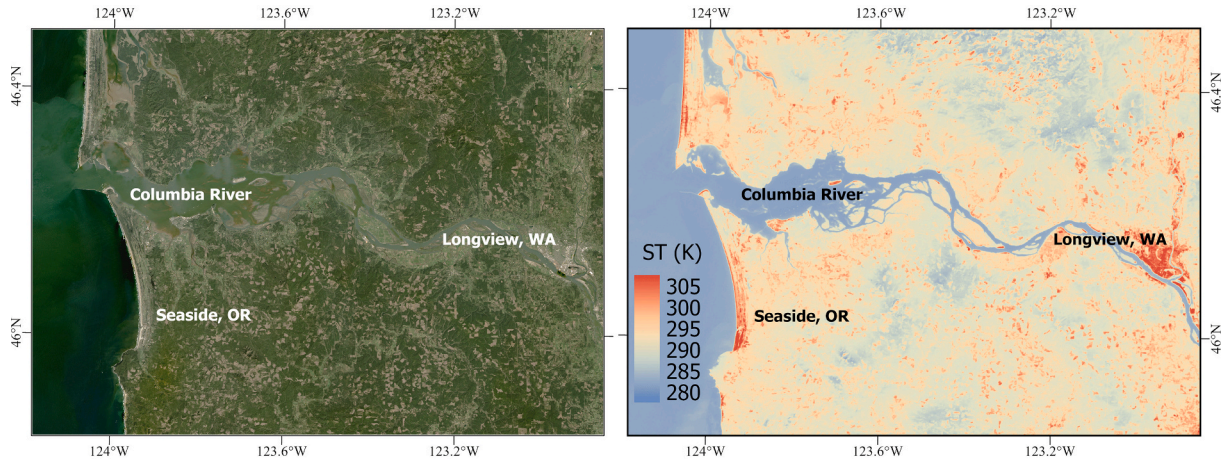


Fig. 1. Example of Landsat 9 observation on April 14, 2025, over World Reference System-2 (WRS-2) path 47/row 28 over the Columbia River in the northwestern United States. The left panel shows Level 2 surface reflectance in true color spectral bands, and the right panel shows the corresponding Collection 2 (C2) Level 2 surface temperature product (Earth Resources Observation and Science (EROS) Center, 2020a). Cooler river and forested areas are contrasted with warmer urban and deforested surfaces.

algorithm (Gerace et al., 2020). Separately, although a stray light correction was implemented operationally by the USGS in 2017 (Gerace and Montanaro, 2017), vicarious calibration indicated the residual stray light effects in Landsat 8 Band 11 (12.0 μm) can still introduce substantial radiometric errors. To minimize the impact of these issues, TIRS Band 10 (10.8 μm) was selected for Landsat 8 single channel algorithm processing, and the same band was used for Landsat 9 to ensure consistency.

The single channel algorithm is based on the radiative transfer equation, where the at-sensor radiance (L^{obs}) for a given wavelength (λ) can be expressed as Eq. (1) (Jiménez-Muñoz and Sobrino, 2003; Jiménez-Muñoz et al., 2008).

$$L_{\lambda}^{\text{obs}} = (L_{\lambda}^T \varepsilon_{\lambda} + (1 - \varepsilon_{\lambda}) L_{\lambda}^{\downarrow}) \tau_{\lambda} + L_{\lambda}^{\uparrow} \quad (1)$$

where L_{λ}^T is radiance emitted by a blackbody at temperature T , ε_{λ} is the emissivity of surface, $L_{\lambda}^{\downarrow}$ is the downwelling radiance, τ_{λ} is atmospheric transmission, and $L_{\lambda}^{\uparrow}$ is the upwelling atmospheric radiance. This equation is solved for L_{λ}^T , and Landsat instrument-specific brightness temperature (BT) lookup tables, pre-computed using Planck's equation, are applied to convert L_{λ}^T to surface temperature. The effective emissivity, atmospheric transmittance, and radiance terms are determined for the Landsat 4–7 Band 6 and Landsat 8–9 Band 10. Landsat 7 ETM+ measures TIR emission in both high and low gain modes (Markham and Helder, 2012), of which the Level 2 surface temperature product may use one or both to provide a single surface temperature quantity that is not negatively affected by radiometric saturation.

3.2. Surface emissivity

Spectral surface emissivity estimates used for processing Landsat 4–9 C2 surface temperature are derived from the ASTER GED Version 003 (GEDv3) (AG100 v003) (Hulley et al., 2015; NASA JPL, 2014), a global, 100-m emissivity dataset of the Earth's non-frozen land surfaces. ASTER GEDv3 contains multi-year averages of derived emissivity from all cloud-free ASTER observations during 2000–2008 for all five TIR spectral bands in the 8–12- μm region (Hulley et al., 2015). The ASTER GEDv3 dataset also includes a mean Normalized Difference Vegetation Index (NDVI) raster layer intended to enable vegetative phenology adjustments of this dataset to changing surface conditions.

For processing Landsat TIR data, ASTER GEDv3 bands are first spectrally adjusted to the Landsat TIR bands. This adjustment is accomplished using the narrow-band emissivity values from ASTER

bands 13 (10.6 μm) and 14 (11.3 μm) along with Landsat sensor-specific regression coefficients developed from laboratory measurements from the ECOSTRESS spectral library (Meerdink et al., 2019). Because ASTER GEDv3 represents multi-year averages, the emissivity values are then adjusted for use with each Landsat observation to account for changes in vegetation phenology and other land surface cover variations. Landsat surface emissivity is modeled as a linear mixture of vegetation and bare soil emission endmembers. This method leverages the NDVI derived from Landsat, along with the mean emissivity and mean NDVI from ASTER GEDv3 (Malakar et al., 2018). Specifically, the Landsat emissivity (ε_L) is estimated as:

$$\varepsilon_L = f_{v,L} \varepsilon_{\text{veg}} + (1 - f_{v,L}) \varepsilon_{\text{bare}} \quad (2)$$

where the $f_{v,L}$ is the vegetation fraction components computed from Landsat, ε_{veg} is the vegetation emissivity, and $\varepsilon_{\text{bare}}$ is the bare soil emissivity derived from ASTER GEDv3. The vegetation emissivity is set to a value between 0.9885 and 0.9899 depending on the TIR band spectral response for each Landsat mission. The bare soil emissivity is derived from the spectrally adjusted ASTER GEDv3 emissivity ($\varepsilon_{\text{ASTER}}$) following Eq. (3).

$$\varepsilon_{\text{bare}} = \frac{\varepsilon_{\text{ASTER}} - \varepsilon_{\text{veg}} f_{v,A}}{1 - f_{v,A}} \quad (3)$$

The vegetation fractions ($f_{v,L}$ and $f_{v,A}$) in Eqs. (2) and (3) are derived from NDVI and normalized using predefined minimum and maximum NDVI values representing bare soil and dense vegetation conditions, respectively, as expressed by:

$$f_{v,i} = 1 - \frac{\text{NDVI}_{\text{max}} - \text{NDVI}_i}{\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}}, i \in \{L, A\} \quad (4)$$

Finally, pixels with a Normalized Difference Snow Index (NDSI) greater than 0.4 are classified as snow and assigned a prescribed band-effective surface emissivity value between 0.9856 and 0.9915 based on TIR band spectral response function of each Landsat mission. For pixels flagged as water in the Landsat quality assessment (QA) raster layer, a mission-specific band-effective water emissivity value close to 0.9880 between 0.9840 to 0.9908 is applied.

3.3. Atmospheric auxiliary data

Solving the single channel equation (Eq. (1)) requires accurate characterization of atmospheric upwelling radiance, downwelling

radiance, and transmittance. These atmospheric parameters are derived from radiative transfer modeling using external atmospheric profiles of specific humidity, geopotential height, and air temperature corresponding to the time and location of Landsat TIR observations. In the C1 provisional surface temperature product, these profiles were obtained from the National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR) (Mesinger et al., 2006), which is limited to North America. To enable global surface temperature generation in C2, USGS evaluated alternative auxiliary atmospheric profile datasets with global coverage (Table 2). Following extensive intercomparisons that considered horizontal resolution, pressure levels, temporal coverage, and data latency, the Goddard Earth Observing System (GEOS) Forward Processing for Instrument Teams (FP-IT) (Rienecker et al., 2008) was adopted as the primary source for acquisitions after January 2000 onward, with the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) (Gelaro et al., 2017) serving as the auxiliary source for acquisitions before 2000, due to the unavailability of FP-IT data before that date. After the release of C2, the NASA Global Modeling and Assimilation Office (GMAO) replaced GEOS FP-IT with GEOS-IT (GEOS for Instrument Teams) (Rienecker et al., 2008), which incorporates an updated GEOS data assimilation system. Consequently, in January 2024, USGS

Table 2

Spatio-temporal resolution, latency, and assimilation system of common global atmospheric reanalysis sources.

Source	Horizontal resolution	Temporal coverage	Temporal resolution	Latency	Assimilation system; URL/DOI
ERA5	$1/4^\circ \times 1/4^\circ$	January 1, 1940 – Present	Hourly	5 days	IFS Cycle 41r2; doi: 10.24381/cds.bd0915c6
MERRA-2	$5/8^\circ \times 1/2^\circ$	January 1, 1980 – Present	3-h	20–50 days	GEOS version 5.12.4; doi: 10.5067/QBZ6MG944HW0
GEOS-FP	$5/16^\circ \times 1/4^\circ$	February 20, 2014 – Present	3-h	12 h	GEOS version is upgraded routinely (GMAO near real-time product); https://gmao.gsfc.nasa.gov/gmao-products/geos-fp/
GEOS FP-IT	$5/8^\circ \times 1/2^\circ$	January 1, 2000 – 2025	3-h	18 h	GEOS version 5.12.4; https://goldsfsl.gesdisc.eosdis.nasa.gov/data/GEOS5/D/FPIT3NPAS/M.5.12.4/
GEOS-IT	$5/8^\circ \times 1/2^\circ$	January 1, 1998 – Present	3-h	18 h	GEOS version 5.29.4; https://goldsfsl.gesdisc.eosdis.nasa.gov/data/GEOSIT/GEOSIT_ASM_I3_L_P4_2.5.29.4/

ECMWF = European Centre for Medium-Range Weather Forecasts, ERA5 = ECMWF Reanalysis version 5, IFS = Integrated Forecasting System, MERRA-2 = Modern-Era Retrospective analysis for Research and Applications, Version 2, GMAO = Global Modeling and Assimilation Office, GEOS = Goddard Earth Observing System, GEOS-FP = GEOS Forward Processing, GEOS FP-IT = GEOS Forward Processing for Instrument Teams, GEOS-IT = GEOS for Instrument Teams, URL = Universal Resource Locator, DOI = Digital Object Identifier.

integrated GEOS-IT into the forward processing of Landsat 8–9 C2 surface temperature products.

The C2 atmospheric auxiliary data from the GEOS assimilation system are provided at 3-h intervals and across 42 atmospheric pressure levels (Rienecker et al., 2008). For each Landsat acquisition, the reanalysis data are temporally and spatially matched to the Landsat scene. First, the reanalysis grid data are horizontally subset to the spatial extent of each Landsat image, with an applied buffer to ensure sufficient coverage at the edges. For each grid point within the subset, atmospheric profiles corresponding to reanalysis times immediately preceding and following the Landsat acquisition time are extracted, converted to the units required by the radiative transfer code, and linearly interpolated to Landsat observation time.

Once spatially subsetted and temporally interpolated, the atmospheric profiles are processed through the MODerate resolution atmospheric TRANsmiSSion (MODTRAN) radiative transfer model Version 5.3.2 (Berk et al., 2006). An external digital elevation model (DEM; Franks et al., 2020) is used to determine the elevation range of the Landsat scene and to select from the 11 surface elevations, ranging from 0 to 9 km. Based on the scene elevation range, the subset of elevations that bound the minimum and maximum scene elevations is identified. MODTRAN is then executed at each reanalysis grid point for each selected elevation (between 2 to 11 elevations) to derive 3-dimensional atmospheric transmittance and upwelling and downwelling radiance quantities. Finally, for each Landsat pixel, the four surrounding reanalysis grid points are identified, and the MODTRAN outputs at those four points are vertically interpolated to the pixel's elevation and then horizontally interpolated to the pixel latitude and longitude using inverse distance weighting.

3.4. Uncertainty estimation

Standard error propagation is applied to compute Landsat surface temperature uncertainty as a function of uncertainties in algorithm input variables (Laraby and Schott, 2018). For this, Eq. (1) is rearranged to solve for L^T as shown in Eq. (5) and standard error propagation is then applied according to Eq. (6).

$$L^T = \frac{L^{obs} - L^1 - (1 - \epsilon)L^1\tau}{\tau\epsilon} \quad (5)$$

Then quantitative uncertainty (S_{L^T}) of Landsat surface temperature is determined via an error propagation model according to Eq. (6):

$$S_{L^T}^2 = S_L^2 + S_A^2 + S_E^2 + S_P + S_U \quad (6)$$

$$= \left(\frac{\partial L^T}{\partial L^{obs}} S_{L^{obs}} \right)^2 + \left(\frac{\partial L^T}{\partial \tau} S_\tau \right)^2 + \left(\frac{\partial L^T}{\partial L^1} S_{L^1} \right)^2 + \left(\frac{\partial L^T}{\partial \epsilon} S_\epsilon \right)^2 + S_P + S_U$$

where, S_L is the uncertainty due to Landsat's at-sensor radiance measurements, S_A is the collective uncertainty in the atmospheric variables (L^1 , L^1 , and τ), S_E is the uncertainty associated with spectral and temporal adjustment of ASTER GEDv3 emissivity, and S_P is correlation uncertainty of the atmospheric variables. To account for underestimations in uncertainty, an “unknown error” term (S_U) was also included. This term was derived based on the findings that single channel algorithm uncertainty increases when clouds are either present or nearby and atmospheric transmission is reduced (Laraby et al., 2016; Schaeffer et al., 2018; Laraby and Schott, 2018). This unknown error term was determined to be inversely correlated with the pixel distance to the nearest cloud detected pixel. In Eq. (6), Landsat TIR measurement uncertainty is assumed to be constant over the mission lifetime, and based on vicarious calibration results, is set to 0.8, 0.4, 0.3, and 0.1 K for Landsat 4–5 TM, Landsat 7 ETM+, Landsat 8 TIRS, and Landsat 9 TIRS,

respectively. The emissivity error (S_e) is computed as the sum of error in ASTER GEDv3 emissivity and spectral adjustment procedure as expressed in Eq. (7).

$$S_e^2 = \left(\sigma_{\text{AGED emis stdev}}^2 + \text{RMSE}_{\text{regression}}^2 + S_{\text{AGED TES algorithm}}^2 \right) \quad (7)$$

where $\sigma_{\text{AGED emis stdev}}$ is the emissivity standard deviations of ASTER GEDv3 bands 13 and 14, $\text{RMSE}_{\text{regression}}$ is the root mean square error (RMSE) of the linear regression fit used in the spectral emissivity adjustment, and $S_{\text{AGED TES algorithm}}$ is the uncertainty of the Temperature and Emissivity Separation (TES) algorithm used to produce the ASTER GEDv3 dataset.

3.5. Validation and consistency

The C1 provisional product for Landsat 5 and 7 was validated against NOAA Surface Radiation Budget (SURFRAD) network and two inland water bodies in situ sites, finding mean biases (RMSEs) of 0.7 K (2.2 K) and 0.9 K (2.3 K) over land, and −0.3 K (0.6 K) and 0.4 K (0.7 K) over water, respectively. These results demonstrated that the operational single-channel algorithm meets the ~1 K accuracy target over water but carries somewhat larger errors over heterogeneous land surfaces (Malakar et al., 2018). A similar evaluation of Landsat C2 surface temperature data for Landsat 8 against JPL automated validation sites showed a mean bias of −0.52 K and precision of 1.0 K over land targets, and −0.24 K mean bias and 0.53 K precision over lake and coastal water bodies (Crawford et al., 2023). Comparisons of Collection 2 against other in situ measurements have been carried out by the remote sensing community at-large. These studies have found some degree of positive bias in Landsat 4–9 surface temperature quantities that include sites over land (+1.0 K, Galve et al., 2022; +0.3 K, Gök et al., 2024), over water (+1.1 K, Vanhellemont et al., 2022; +2.4 K, Dyba et al., 2022; +0.5–0.9 K, Herrick et al., 2023), and over areas of intermittent inundation (+0.4 K, Niclòs et al., 2023). Variability among these published results is to be expected and possibly arises from factors such as site-specific differences in average temperature, surface emissivity, atmospheric conditions, differences in in situ data collection techniques and instrumentation, the subset of the Landsat record being examined, or perhaps differences in data reduction approaches such as estimating water skin surface temperatures from measurements taken at depth (Wilson et al., 2013). Overall, these studies demonstrate good agreement between Landsat C2 surface temperature and a variety of in situ measurements, with biases that are generally small. However, mismatches between estimated and actual emissivity quantities have been shown to be a major source of disagreement between satellite-derived land surface temperature and in situ measurements (Galve et al., 2022).

In addition to validation against in situ ground measurements, internal consistency between Landsat missions across the data record to support multidecadal surface temperature variability and trend analyses are important considerations because the long observational record is a major strength of the Landsat program. Many studies utilizing the Landsat Collection 2 products are not formal validation efforts, but research results encompassing both the challenges and successes of working with Landsat data are relevant for Collection 3 development, including use of these data in long-term trending analyses. As examples, previous studies have found that the Landsat C2 surface temperature products can be used to investigate lake surface temperature dynamics (Herrick et al., 2023) and long-term trends over mountainous terrain (Gök et al., 2024), although the two studies both noted that careful data handling is required to interpret surface temperature trends reliably. These authors found that additional data filtration steps were required to remove outliers and spurious trends, which are often due to issues not directly tied to the atmospheric correction algorithm, such as the presence of undetected clouds or off-nominal satellite overpass times.

3.6. Collection 2 known limitations

Several known issues affecting the Landsat C2 surface temperature product have been documented by the USGS (USGS Landsat Missions, 2022). These issues include spatial artifacts that are visually apparent in the surface temperature products and reflect C2 processing shortcomings that are not readily discoverable in quantitative validation alone (Table 3). Several of these issues have also been reported in the literature. For example, spatial discontinuities in the form of a checkerboard pattern have been observed by Vanhellemont et al. (2022) and Xie et al. (2024), and temporal analyses have found cloud omission errors in the pixel QA band (Gök et al., 2024; Herrick et al., 2023). Issues related to undetected clouds pertain to errors from cloud masking applied to Landsat Level 1 data which, although relevant for Landsat Collection 3 development more generally, are beyond the scope of this work. The other issues and known limitations in the C2 surface temperature product are described in following subsections, and Section 4 details how the USGS will update and improve in C3.

3.6.1. Data gaps and mischaracterized surface emissivity

The primary limitation of the Landsat C2 surface temperature product over land arises from mischaracterization of surface emissivity. Three specific issues contribute to this limitation. First, NoData values occur in the C2 surface temperature product where emissivity data are unavailable. These missing values originate from either NoData gaps in the ASTER GEDv3 dataset or masking of unphysical emissivity values within the Landsat emissivity estimation workflow. The ASTER GEDv3 represents a multi-year average of all cloud-free observations of Earth's non-frozen land surfaces during 2000–2008. Equatorial regions with persistent cloud cover (e.g., portions of Amazon rainforest and Indonesia) or high latitude areas with perennial snow/ice cover (e.g., Antarctica and northern Eurasia) lack valid ASTER emissivity data, resulting in persistent NoData in subsequent Landsat surface temperature retrievals (Fig. 2). In addition, after applying spectral and temporal vegetation phenology adjustments, the Landsat emissivity retrieval workflow imposes a lower threshold on adjusted Landsat emissivity ($\epsilon < 0.6$) for natural land surfaces; pixels below this threshold are also set to NoData, resulting in more occurrences in the Landsat C2 surface temperature product.

Second, an anomaly was identified in the temporal vegetation emissivity adjustment procedure described in Section 3.2. In C2, pixels classified as ASTER bare soil in the ASTER GEDv3 (ASTER NDVI < 0.5) are overwritten and assigned the bare soil emissivity value, as defined in Eq. (3). This bare soil override was introduced to prevent NDVI-based emissivity adjustments in desert regions, where NDVI and emissivity are weakly correlated. However, this approach leads to more uncertainty in Landsat emissivity estimation, and consequently, surface temperature retrieval for pixels that have undergone substantial vegetation cover changes between the ASTER GEDv3 observation period (2000–2008) and the Landsat acquisition date. This effect is particularly pronounced in areas with strong temperature and/or emissivity gradients such as arid and semi-arid lands that have been converted to irrigated agriculture fields. In addition, in the C2 implementation of Eq. (2), the vegetation spectral index (NDVI) is used directly in place of vegetation fraction ($f_{v,l}$). Because NDVI is not standardized between ASTER and Landsat observations, this substitution can result in inconsistent vegetation adjustments and additional emissivity error.

Third, the adjusted Landsat emissivity often exhibits a blocky spatial appearance. This artifact arises from two factors. First, ASTER GEDv3 is provided at 100-m resolution and is resampled to the Landsat 30-m grid using nearest-neighbor interpolation. Such a resampling method is not well suited for continuous variables such as emissivity. Second, in ASTER GEDv3, NDVI derived from the visible and near infrared (VNIR) instrument is co-registered to TIR instrument data that has limited geolocation accuracy in some locations (Abrams et al., 2002). This misregistration introduces a spatial offset between ASTER NDVI and

Table 3

Summary of the main Collection 2 (C2) limitations and Collection 3 (C3) improvements for Landsat surface temperature and emissivity (the section describing the change is shown in parenthesis where applicable).

Processing step	C2 limitation	C3 enhancement	Product impact
Atmospheric correction algorithm	Single channel for Landsat 8 and 9 does not use TIRS Band 11 (Section 3.1)	Evaluation of generalized split window for Landsat 8 and 9 (Section 4.3)	Improved surface temperature retrieval in certain atmospheric and surface conditions
Surface emissivity estimation	Gaps in ASTER GEDv3 (Section 3.6.1)	Use of CAMEL V003 to fill ASTER GEDv3 gaps (Section 4.1.1)	True global coverage, including polar regions and small islands, as well as nighttime data coverage
Surface emissivity estimation	Limited vegetation phenology adjustment for temporal variations in surface emissivity (Section 3.6.1)	Implementation of corrected vegetation phenology adjustment (Section 4.1.2)	More accurate Landsat surface emissivity, particularly where land cover type changed considerably since the 2000–2008 ASTER GEDv3 period
Surface emissivity estimation	Mischaracterized surface emissivity due to nearest neighbor resampling of ASTER GEDv3 (Section 3.6.2)	Use of cubic convolution resampling method (Section 4.1.2)	More realistic representation of spatial variation at Landsat GSD
Radiative transfer code for single channel	MODTRAN v 5.3.2 (Section 3.3)	RTTOV v 13.1 or higher (Section 4.2.1)	Faster processing and use of open-source software
Characterization of atmospheric parameters	Quadratic interpolation of reanalysis grid points with topographic imprinting features (Section 3.6.2)	Use of bicubic interpolation of auxiliary data sources (Section 4.2.2)	Smooth mesoscale variations of atmospheric parameters as expected, and no more blocky artifacts
Standard error propagation	Overestimation of uncertainty due to distance to cloud method (Section 3.6.3)	Removal of distance to cloud and refinement of uncertainties associated with atmospheric state (Section 4.2.3)	Reduced overestimation of uncertainty in the presence of false positive cloud detected pixels
Processing of Level 2 products from VSWIR and TIR observations	Coupled processing of Level 2 surface reflectance and surface temperature (Section 3.6.4)	Decoupling of surface temperature and emissivity from surface reflectance processing (Section 4.4)	Enables polar, twilight, and nighttime coverage of Level 2 surface temperature and emissivity
Surface temperature upper limit	373 K cap	Extension of cap to 500 K (Section 4.4)	Improved mapping of wildfire and lava temperatures
External DEM source	Combination of NASADEM, ArcticDEM, Alaska NED, CDEM, and SNF DEM	Use of COPDEM (Section 4.4)	Improved spatial consistency
External atmospheric reanalysis source	MERRA-2, GEOS FP-IT, GEOS-IT	MERRA-2 or MERRA-3 (if available) (Section 4.4)	Improved temporal consistency

C2 = Collection 2, C3 = Collection 3, TIRS = Thermal Infrared Sensor, ASTER = Advanced Spaceborne Thermal Emission and Reflection Radiometer, GED = Global Emissivity Dataset, CAMEL = Combined ASTER and Moderate Resolution Imaging Spectroradiometer (MODIS) Emissivity for Land, MODTRAN = Moderate resolution atmospheric TRANsmission, RTTOV = Radiative Transfer for Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS), VSWIR = Visible to Short Wave Infrared, TIR = Thermal Infrared, DEM = Digital Elevation Model, NASADEM = National Aeronautics and Space Administration DEM, NED = National Elevation Dataset, CDEM = Canadian DEM, SNF = Sweden Norway Finland, COPDEM = Copernicus DEM, MERRA-2 = Modern-Era Retrospective analysis for Research and Applications, Version 2, MERRA-3 = Modern-Era Retrospective analysis for Research and Applications, Version 3, GEOS = Goddard Earth Observing System, FP-IT = Forward Processing for Instrument Teams, GEOS-IT = GEOS for Instrument Teams.

Landsat NDVI and further accentuates the blocky pattern in the adjusted emissivity raster layer.

3.6.2. Suboptimal interpolation of atmospheric profiles

The C2 auxiliary atmospheric profile datasets from MERRA-2, GEOS FP-IT, and GEOS-IT, described in [Section 3.3](#), have a horizontal resolution of $5/8^\circ \times 1/2^\circ$ (roughly 50 km). After the atmospheric profiles are spatially subset and temporally interpolated to the time of Landsat acquisition, MODTRAN Version 5.3.2 is executed and the resulting atmospheric transmittance, upwelling, and downwelling radiance are interpolated to the Landsat pixel elevation and corresponding 30-m sampling resolution. The horizontal interpolation is performed using the four surrounding reanalysis grid points. This approach introduces visible seamlines (i.e., checkerboard patterns) in the atmospheric parameter fields which can propagate into the final surface temperature. In addition, the vertical interpolation to the pixel elevation causes fine-scale surface topography variations to become imprinted in the final atmospheric layers. While atmospheric conditions are expected to vary with elevation, radiative transfer parameters generally exhibit smooth mesoscale gradients over kilometers scale rather than variability at the scale of surface topography ([Jedlovec, 1990](#); [Rozenstein et al., 2014](#); [Thompson et al., 2021](#)).

3.6.3. Overestimation of surface temperature uncertainty

The uncertainty estimation approach described in [Section 3.4](#) relies on the distance to the nearest cloud detected pixel. While this method generally mitigates uncertainty underestimation inherent in the standard error propagation model, it introduces a dependency on cloud detection algorithm accuracy. Errors in cloud detection produce unrealistic uncertainty estimates, as well as spatial artifacts. For example, false-positive cloud detections result in artificially elevated surface temperature uncertainty values in the C2 product.

3.6.4. Landsat Level 2 product processing flow

Production of Landsat C2 atmospherically corrected surface temperature is bundled with surface reflectance within the USGS Landsat Product Generation System (LPGS; <https://www.usgs.gov/landsat-missions/landsat-processing-systems>) with both geophysical products being delivered together as a Level 2 product package. This data-processing architecture requires that all atmospheric and surface auxiliary inputs needed for both Level 2 surface reflectance and surface temperature product generation be available for each Landsat scene. While a surface reflectance-only product may be generated for scenes lacking auxiliary surface emissivity (e.g., over Antarctica), C2 does not support generation of surface temperature-only products. This limitation is partly attributable to the Landsat surface emissivity retrieval, which relies on vegetation and snow spectral indices derived from atmospherically corrected surface reflectance data for phenological emissivity adjustments as described in [Section 3.2](#).

This product generation dependency becomes problematic when surface temperature can be retrieved under illumination conditions for which no operational Landsat C2 Level 2 surface reflectance product is

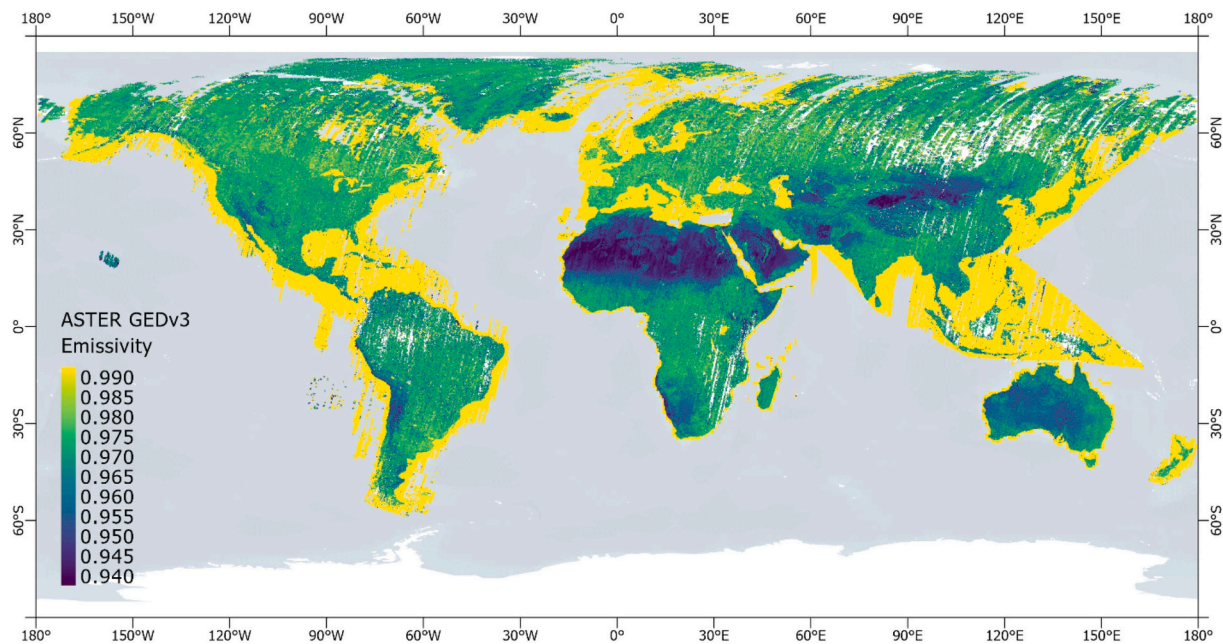


Fig. 2. Geographic coverage of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset Version 003 (GEDv3). The dataset represents mean emissivity and was generated from cloud-free ASTER observations of Earth's non-frozen land surfaces collected during 2000–2008. Areas lacking mean emissivity due to persistent cloud cover (e.g., portions of the Amazon rainforest) or permanent snow and ice cover (e.g., Antarctica and northern Eurasia) are shown in white.

generated. In C2, Landsat Level 2 surface reflectance is not produced for scenes with a scene-center solar zenith angle (SZA) greater than 76° . Under high SZA conditions, longer downwelling atmospheric path lengths result in a loss of atmospheric correction fidelity under the plane-parallel assumptions inherent to the radiative transfer model used for visible-to-shortwave infrared (VSWIR) atmospheric compensation (Kotchenova et al., 2006). In principle, an accurate atmospheric correction of emitted TIR radiance can still be performed provided an acceptable substitute for emissivity correction can be identified. Removing this C2 Level 2 production constraint would allow surface temperature retrieval for high-latitude Arctic and Antarctic regions, and for nighttime acquisitions to be included in C3 Level 2 surface temperature emissivity processing.

4. Collection 3 improvements

Since the adoption of collection-based processing for the Landsat data record in 2016, the USGS has completed two major reprocessing campaigns. Landsat C1 was available from September 2016 through December 2022 and provided consistently calibrated Level 1 products from Landsat 1–8. The C1 Level 2 atmospherically corrected surface reflectance and surface temperature products were generated for Landsat 4–8 as part of U.S. ARD (Dwyer et al., 2018). C2 was released in December 2020 and is operational at the time of writing (2026). Collection 2 built on the success and framework of C1 by incorporating Landsat 9 data after its launch in September 2021. For C2, the USGS improved Landsat absolute geolocation accuracy, enabled direct access through commercial cloud services, and released for the first time, an inventory of globally distributed Level 2 surface reflectance and surface temperature products (Crawford et al., 2023). The USGS is planning for release of Landsat C3 no earlier than December 2028. C3 presents an opportunity to adopt and integrate advances in Landsat remote sensing algorithms and processing architectures since the release of C2, address known issues with C2 products, and establish a pathway for incorporating Landsat 10 science data in the early 2030s (Roy et al., 2026).

4.1. Emissivity estimation enhancements

As noted in Section 3.6.1, three specific issues affect the surface emissivity retrieval in the Landsat C2 surface temperature product: data gaps in ASTER GEDv3, anomalies in vegetation phenology adjustment, and the blocky spatial pattern in the estimated Landsat emissivity. For each of these issues, mitigation solutions have been developed for C3 reprocessing and are described in following subsections.

4.1.1. Emissivity gap filling

To address data gaps in the ASTER GEDv3 that trace to persistent cloud or snow/ice conditions, the Combined ASTER and MODIS Emissivity over Land (CAMEL) Version 003 (V003) dataset (CAM5K30EM V003; Borbas et al., 2023) has been adopted as a gap filling emissivity source for C3. The CAMEL V003 dataset provides monthly mean surface emissivity at 0.05° (~ 5 -km) spatial resolution across 13 spectral bands spanning from 3.6 to $14.3 \mu\text{m}$. The dataset is generated by merging the MODIS baseline-fit emissivity database (MODBF) with the monthly mean ASTER GED Version 004 (AG5KMMOH) (Borbas et al., 2018; Feltz et al., 2018).

The CAMEL Version 003 dataset provides coverage of global land-masses and near-shore coastal regions from 2000 through 2023. For Landsat emissivity estimation in C3, the ASTER GEDv3 dataset remains the primary emissivity input source due to its higher spatial resolution, which is closer to the native spatial resolution of Landsat TIR observations. For locations where ASTER GEDv3 contains NoData values, emissivity is sourced from CAMEL V003. In this scenario, the CAMEL emissivity values are spectrally adjusted to Landsat TIR band spectral response function using linear regression and precomputed coefficients. Because CAMEL V003 provides monthly climatological surface emissivity, a temporal vegetation phenology adjustment is not required. Together, ASTER GEDv3 and CAMEL V003 enable complete global coverage for C3 Landsat surface temperature and emissivity products, including polar regions. Fig. 3 illustrates an example of Landsat 9 surface emissivity retrieved from ASTER GEDv3 (C2 product), ASTER GED gap-filled with CAMEL V003 (future C3 product), and the corresponding retrieved surface temperature. The provisional C3 product provides

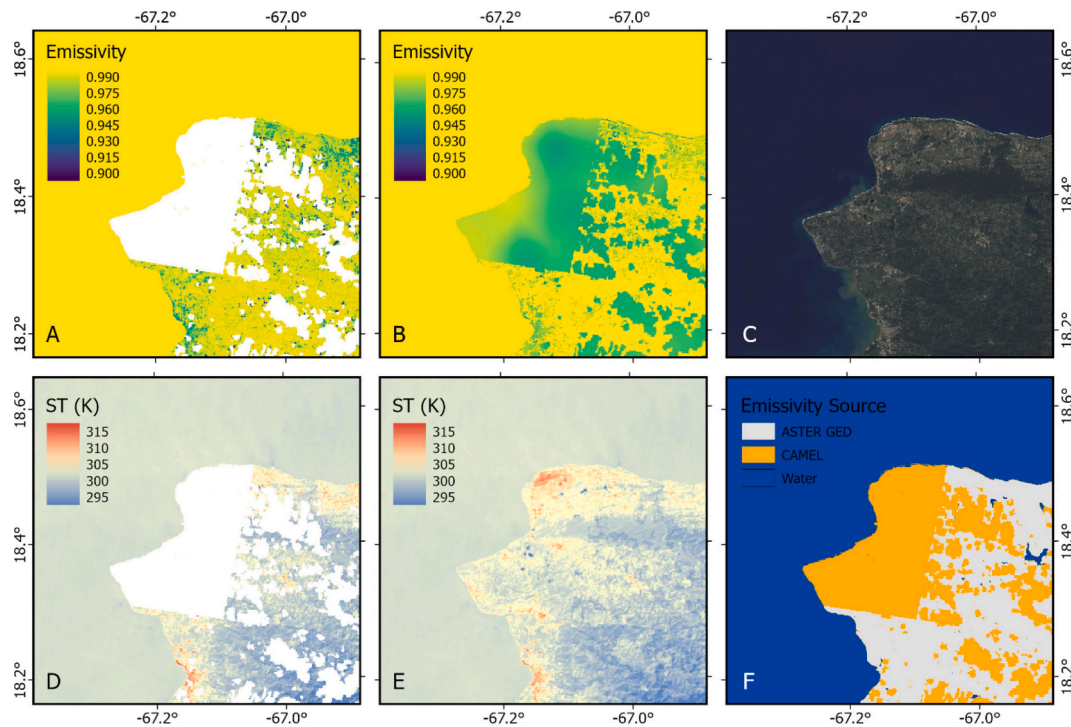


Fig. 3. Landsat 8 imagery acquired on January 6, 2025, over World Reference System-2 (WRS-2) path 5/row 47. Emissivity data are shown for Collection 2 (C2) (panel A) and the gap filled Collection 3 (C3) emissivity (panel B). The surface temperature output using the respective emissivity layers are shown in panels D and E. White indicates NoData pixels in panels A and D. Panel C provides a true color visual for context, and panel F is the emissivity source layer providing per-pixel context to where emissivity value was derived from.

valid surface temperature data values in the area of the scene where there is no valid data in the C2 product. A per-pixel emissivity source layer similar to panel F will be included in the C3 Level 2 surface temperature and emissivity product (panel F in Fig. 3). This layer will document whether ASTER GEDv3, CAMEL, or prescribed snow or water emissivity values were used to estimate Landsat spectral surface emissivity, thereby improving product transparency and traceability.

To reduce visible seamlines introduced by combining ASTER GEDv3 and CAMEL emissivity layers, a scene-level decision rule was implemented whereby CAMEL emissivity is used exclusively when ASTER GEDv3 coverage within a Landsat image falls below 75%. Under this condition, CAMEL provides emissivity for the entire scene, rather than being used solely for NoData gap filling. Fig. 4 illustrates a comparison

of Landsat surface emissivity derived from ASTER GEDv3 only (C2 product), ASTER GEDv3 with CAMEL gap filling, and full-scene CAMEL emissivity (future C3 product).

For Landsat acquisitions between March 2000 and December 2023, the CAMEL V003 monthly emissivity product corresponding to the month of Landsat image acquisition will be used as a source for filling emissivity gaps. For Landsat 4–7 acquisitions prior to March 2000, the earliest available CAMEL V003 product from 2000 or 2001 that matches the month of Landsat image acquisition will be used as a surrogate. Because the CAMEL V003 does not currently extend beyond January 2024, processing Landsat 8–9 data beyond this date relies on the CAMEL data from the corresponding month of 2023. A new NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs)

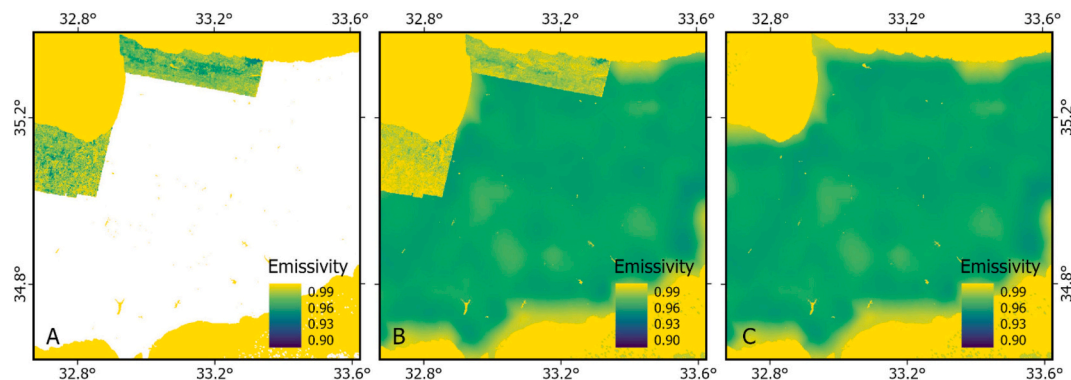


Fig. 4. Example of Landsat surface emissivity derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset Version 003 (GEDv3) in Collection 2 (C2) (panel A; white indicates NoData), ASTER GEDv3 combined with monthly Combined ASTER and MODIS Emissivity over Land (CAMEL) V003 (panel B), and CAMEL V003 only emissivity (panel C). This Landsat 9 image of Cyprus was acquired on September 5, 2022, over World Reference System-2 (WRS-2) path 176/row 36.

emissivity dataset is anticipated prior to C3 reprocessing to expand the temporal coverage of CAMEL V003 beyond January 2024. This forthcoming dataset will likely incorporate MODIS observations from recent years and then transition to VIIRS-based emissivity following the planned decommissioning of the Terra and Aqua missions in late 2026 and early 2027 (NASA Earthdata, 2025).

In the context of future Landsat C3 surface temperature and emissivity forward processing, the surface emissivity data latency from CAMEL V003 will modulate the Landsat 8 and 9 Level 2 product generation timeline. Given the monthly cadence of CAMEL V003, if emissivity data for the month of a Landsat image acquisition are unavailable, the processing logic will select the nearest available monthly product within a ± 2 -month window. If no suitable match is found within that timeframe, the search window will be expanded to one year, and the emissivity product from the same month (± 2 months) of the previous year is used to preserve seasonal consistency associated with land surface phenology. If no valid emissivity data are available, the search is iteratively expanded to additional years while maintaining the same seasonal window until a suitable product is identified.

ECOSTRESS observations from the ISS (Hook and Hulley, 2019) represent another potential emissivity source for filling NoData gaps in ASTER GEDv3 at midlatitudes (55°S to 55°N). With five ASTER-like TIR spectral bands and a spatial resolution of 70 m, ECOSTRESS has the potential to enhance gap filling and resolve the fine-scale spatial heterogeneity in surface emissivity, particularly in tropical regions where ASTER NoData values are more prevalent. The science quality ECOSTRESS observations could be used to construct a multi-year average emissivity dataset analogous to ASTER GEDv3, prior to integration into the Landsat surface temperature and emissivity processing framework. With the planned decommissioning of the ISS in the 2030s (NASA, 2024), a medium resolution ECOSTRESS-based emissivity dataset could provide an important bridge between the ASTER GEDv3 and

upcoming or planned TIR missions (e.g., TRISHNA, LSTM and EAGLE-TIR) until Landsat 10 is launched in the early 2030s.

4.1.2. Enhanced resampling and temporal adjustment of ASTER GEDv3

The blocky appearance of Landsat C2 surface emissivity described in Section 3.6.1 results in part from resampling the 100-m ASTER GEDv3 to 30 m using nearest neighbor interpolation. To address this, the resampling method has been replaced with cubic convolution, which is generally preferred for continuous quantities (Keys, 1981). In addition, the Landsat emissivity retrieval has been revised to address the mischaracterization issues described in Section 3.6.1 and to align with the temporal vegetation phenology adjustment procedure described in Eq. (2) through Eq. (4) (Section 3.2). The revised implementation for C3 uses the vegetation fraction ($f_{v,L}$), as shown in Eq. (2), instead of NDVI-based formulation used in C2. The bare soil emissivity also no longer overrides pixels with ASTER NDVI < 0.5 . Instead, to account for instability in bare soil emissivity derived from ASTER GEDv3, a prescribed bare soil emissivity is substituted for ϵ_{bare} (Eq. (3)) at locations where the ASTER vegetation fraction ($f_{v,A}$) exceeds 0.8. To mitigate spatial heterogeneity introduced by this substitution, a 5×5 smoothing kernel is applied to the resulting ϵ_{bare} prior to its use in Eq. (1). Fig. 5 shows the combined effect of cubic convolution resampling of 100-m ASTER GEDv3 and the revised C3 temporal vegetation phenology adjustment when compared to the C2 Landsat emissivity layer over center pivot agricultural fields.

4.1.3. Prescribed emissivity assignment

As described in Section 3.2, the C2 surface temperature processing uses an NDSI > 0.4 threshold to identify snow and ice pixel conditions and assigns prescribed emissivity quantities. This technique produces commission errors where snow-flagged pixels appear in unexpected locations. To address this issue for the C3 STE product, the snow/ice flag

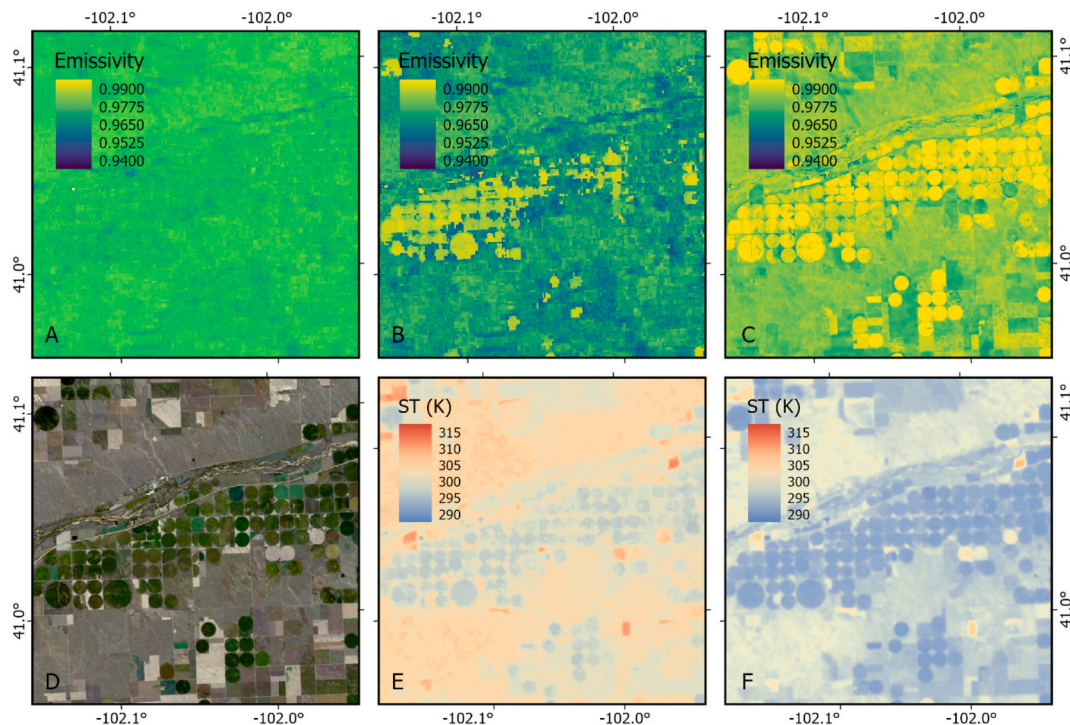


Fig. 5. Example of the original Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset Version 003 (GEDv3) at ~ 100 m resolution (panel A), Landsat Collection 2 (C2) surface emissivity adjusted from ASTER GEDv3 using nearest neighbor resampling (panel B), Landsat Collection 3 (C3) surface emissivity with cubic convolution resampling and phenology adjustments using vegetation fraction (panel C), true color surface reflectance (panel D), the retrieved C2 ST (panel E), and prototype C3 ST (panel F). The figure demonstrates that the C3 outputs now show more realistic spatial patterns of surface emissivity at the time of Landsat overpass. This Landsat 9 image of center pivot agricultural fields in Colorado was acquired on September 28, 2025, over World Reference System-2 (WRS-2) path 32/row 31.

from the Landsat per-pixel QA layer will be used to identify snow and ice pixel conditions and assign the prescribed emissivity quantities described earlier. The CFMask 3.3 cloud masking algorithm used in Landsat C2 processing applies more rigorous spectral tests for snow and ice detection (Foga et al., 2017; Zhu et al., 2015). These spectral tests incorporate thresholds in the Near Infrared (NIR), Green, and TIR spectral bands in addition to NDSI, which produce fewer false positives. Adoption of a QA-based snow mask is therefore expected to improve the accuracy of the prescribed emissivity and would further inherit any additional benefits from improvements to the pixel QA layer in-development for Collection 3.

4.2. Single channel algorithm implementation enhancements

4.2.1. Radiative transfer code

In C2, the single channel algorithm uses MODTRAN Version 5.3.2 radiative transfer code to compute the atmospheric transmittance, upwelling radiance, and downwelling radiance (Berk et al., 2006). For C3, MODTRAN will be replaced with the open-source Radiative Transfer for Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) (RTTOV) Version 13 or newer (Saunders et al., 2018). RTTOV employs a parameterized radiative transfer approach, using pre-computed coefficients and regression-based predictors to estimate atmospheric transmittances rather than performing detailed line-by-line spectral integration for each simulation. The RTTOV approach substantially reduces the computational burden of the radiative transfer execution step within the single channel algorithm while maintaining computational accuracy, making it well suited for large scale operational processing of the C3 surface temperature and emissivity products.

In addition to specific humidity and air temperature variable inputs used for the MODTRAN radiative transfer model, RTTOV also requires surface pressure, all of which are assimilated from external atmospheric auxiliary datasets. A comparative analysis of surface temperature retrieved using MODTRAN and RTTOV is presented in Section 5.2.

Under cloud-free conditions, the mean difference between the output of the two radiative transfer codes is 0.1 K with a standard deviation of 0.34 K, with 95% of pixels showing differences less than 0.83 K. The largest discrepancies occur in regions with steep topography or strong atmospheric humidity gradients. However, even in these areas, surface temperature differences rarely exceed 1 K.

4.2.2. Atmospheric auxiliary data

In Section 3.3, we described how the auxiliary atmospheric profiles are spatially subset and temporally interpolated to each Landsat image. In C2, the MODTRAN code is executed at multiple elevations, followed by interpolation to the Landsat pixel elevation and geographic location. In C3, RTTOV will be executed only at the ground elevation determined from an external DEM, thereby eliminating the vertical interpolation step. Atmospheric parameters computed at the reanalysis grid points are horizontally interpolated to 30 m using bicubic interpolation to produce smoothly varying mesoscale atmospheric fields. This refined approach for C3 resolves both the seamline artifacts (checkerboard pattern) present in C2 and the presence of topographic features at Landsat pixel scale. Fig. 6 illustrates an example of atmospheric state computed from C2 and the corresponding outputs from the revised processing approach for C3.

4.2.3. Uncertainty estimation enhancements

Further assessment of the standard error propagation model used for C2 single channel uncertainty retrieval (Section 3.4) showed that the “unknown error” term (S_U) in Eq. (6) primarily compensated for an underestimation of atmospheric uncertainty (S_A). Uncertainties in atmospheric transmittance, upwelling radiance, and downwelling radiance is computed using a second-order polynomial regression, with coefficients derived from Monte Carlo simulations. Updated regression coefficients for C3 address the C2 underestimation and eliminate the need for the distance-to-cloud based adjustment included in the C2 single channel error propagation model. Fig. 7 compares single channel

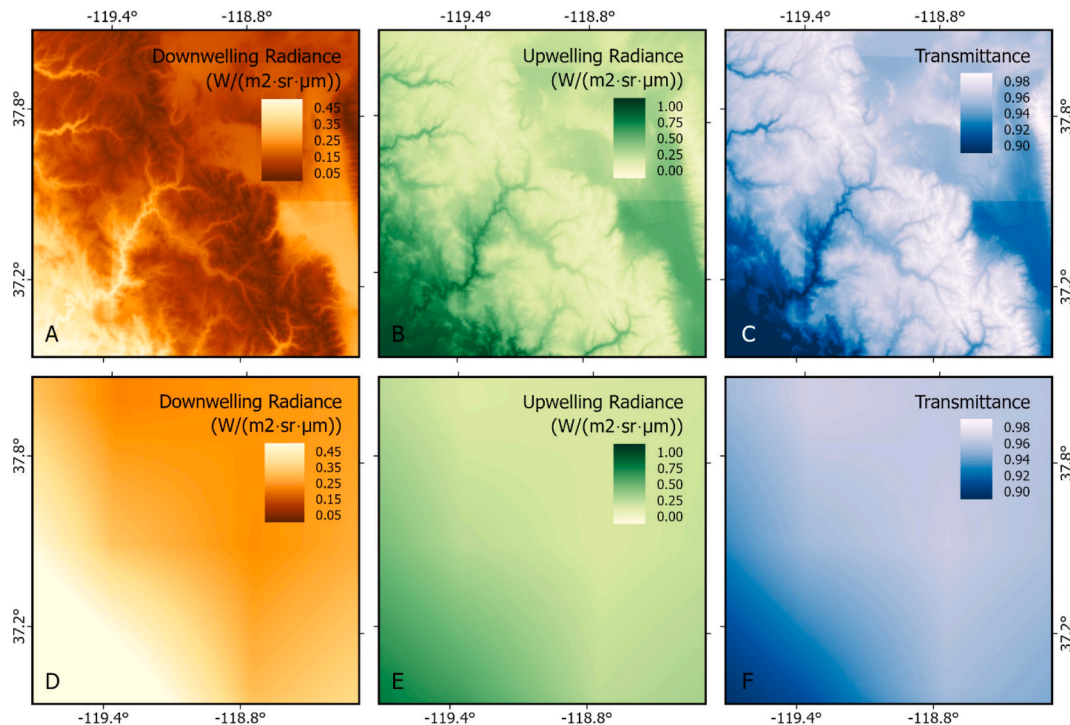


Fig. 6. Example of downwelling radiance (panel A), upwelling radiance (panel B), and atmospheric transmittance (panel C) computed using the Collection 2 (C2) single channel algorithm. Checkerboard patterns and topographic features are visible in the C2 atmospheric state whereas for Collection 3 (C3), the atmospheric fields smoothly vary (three lower corresponding panels D-F). The atmospheric state examples displayed here correspond to a Landsat 7 image of the eastern Sierra Nevada, California acquired on September 27, 1999, over World Reference System-2 (WRS-2) path 42/row 34.

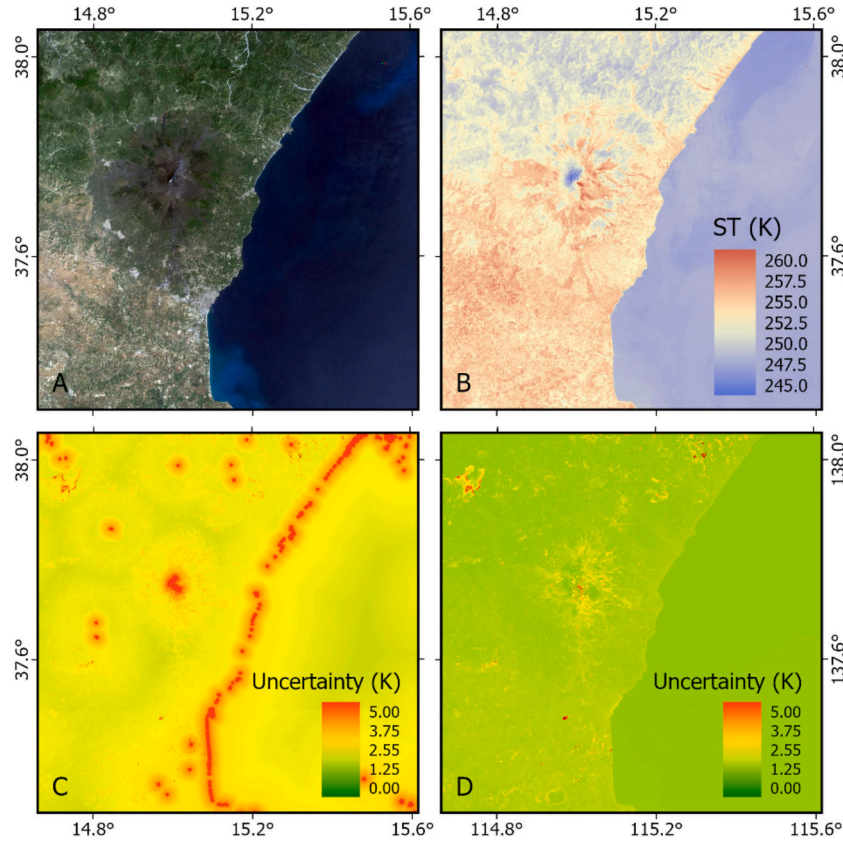


Fig. 7. Example of Landsat 5 true color surface reflectance (panel A), single channel surface temperature (panel B), Collection 2 (C2) single channel algorithm uncertainty with distance to cloud used in standard error propagation model (panel C), and revised surface temperature uncertainty model for Collection 3 (C3) single channel algorithm using updated atmospheric coefficients (panel D). Elevated uncertainty values in panel C are artifacts caused by false positive cloud detection over bright targets (e.g. coastal sand) rather than actual increases in surface temperature uncertainty. The revised C3 model mitigates these artifacts. This Landsat 5 image of Mt. Etna, Italy was acquired on May 19, 1984, over World Reference System-2 (WRS-2) path 188/row 34.

surface temperature uncertainty estimates from C2 and the revised C3 model and demonstrates improved and more consistent results. Further research and development is underway to assess the feasibility of potentially incorporating total precipitable water (TPW) in the uncertainty estimation model to improve further improve accuracy and consistency.

4.3. Evaluation of split window for Landsat 8 and 9

The split window algorithm for retrieving surface temperature was originally developed for the Advanced Very High Resolution Radiometer (AVHRR) to correct for atmospheric absorption using differential absorption in adjacent TIR bands (Becker and Li, 1990). This technique was subsequently generalized to support a broader class of dual-channel TIR instruments by explicitly accounting for surface emissivity effects and incorporating sensor-specific coefficients (Wan and Dozier, 1996). Further refinements were introduced for MODIS, including improved parameterizations to enhance performance over bare soils and heterogeneous land surfaces (Wan, 2008). Following the launch of Landsat 8 in 2013, the generalized split window formulation was adapted to the TIR spectral response functions and viewing geometry of Landsat 8 TIRS, enabling improved land surface temperature retrievals for Landsat observations (Jimenez-Munoz et al., 2014; Gerace et al., 2020).

In the generalized split window algorithm, atmospheric correction is achieved using the differential and average emissivity of two adjacent TIR spectral bands together with sensor-specific coefficients derived from radiative transfer simulations. The surface temperature (ST) is expressed as:

$$ST = b_0 + \left(b_1 + b_2 \frac{1 - \varepsilon}{\varepsilon} + b_3 \frac{\Delta \varepsilon}{\varepsilon^2} \right) \frac{T_i + T_j}{2} + \left(b_4 + b_5 \frac{1 - \varepsilon}{\varepsilon} + b_6 \frac{\Delta \varepsilon}{\varepsilon^2} \right) \frac{T_i - T_j}{2} + b_6 (T_i - T_j)^2 \quad (8)$$

where ST is the surface temperature, b_i are the regression coefficients derived from simulated data, T_i and T_j are the corresponding Top of Atmosphere (TOA) BT, and ε and $\Delta \varepsilon$ are the mean and difference in surface emissivity between the two adjacent TIR spectral bands. The surface emissivity retrieval for each TIR spectral band mirrors the phenology adjustment technique described in Section 3.2 for single channel algorithm, with the enhancements outlined in Section 4.1 for C3.

Recently, an uncertainty estimation method for the split window algorithm has been developed by Hassanzadeh et al. (2025). This method is primarily based on standard error propagation following Guide to the Expression of Uncertainty in Measurement (GUM) (ISO and OIML, 1993), in which the contribution of each term in Eq. (8) is quantified while accounting for correlations among variables. Testing and validation of this split window uncertainty method indicated that uncertainty was underestimated when compared to validation data, and thus, an additional numerical term was incorporated in the split window uncertainty to explicitly account for uncertainty in atmospheric water vapor, as described in Eq. (9).

$$S_{ST}^2 = S_T^2 + S_E^2 + S_P + S_{TPW}^2 \quad (9)$$

$$= \left(\frac{\partial ST}{\partial T_{i/j}} S_{T_{i/j}} \right)^2 + \left(\frac{\partial ST}{\partial \epsilon_{i/j}} S_{\epsilon_{i/j}} \right)^2 + \left(2\rho(T)_{i/j} \frac{\partial LST}{\partial T_i} \frac{\partial LST}{\partial T_j} S_{T_i} S_{T_j} \right. \\ \left. + 2\rho(\epsilon)_{i/j} \frac{\partial LST}{\partial \epsilon_i} \frac{\partial LST}{\partial \epsilon_j} S_{\epsilon_i} S_{\epsilon_j} \right) + S_{TPW}^2$$

where,

$$S_{TPW}^2 = S_{Model}^2 + S_{Fit}^2 + \left(\frac{\partial Model}{\partial TPW_{est}} S_{TPW_{est}} \right)^2 \quad (10)$$

In Eq. (9), $S_{T_{i/j}}$ denotes the sensor noise in the apparent temperature of each Landsat 8 and 9 TIRS spectral band, expressed as the absolute radiometric accuracy; $S_{\epsilon_{i/j}}$ is the emissivity uncertainty for each spectral band; and S_{TPW} represents the uncertainty in the split window algorithm error with respect to varying levels of total precipitable water (TPW). A data-driven model has been developed to estimate the intermediate TPW and corresponding S_{TPW}^2 by mapping the Landsat 8 and 9 TIRS Bands 10 and 11 radiances to TPW, using the Aerosol Robotic Network (AERONET) observations as the reference dataset (Hassanzadeh et al., 2025). Fig. 8 presents an example of the TPW field and the corresponding split window uncertainty from the standard error propagation method planned for the C3 implementation. Operationally, TPW and its associated uncertainty can also be estimated from external atmospheric state datasets, such as the reanalysis products described in Section 4.2.2.

To reduce aliasing artifacts caused by Landsat 8–9 TIRS Band 10 and Band 11 sampling out-of-phase due to their physical offset on the focal plane, a low pass filter using a 5×5 convolution kernel is applied to the

Level 1 TOA BT data prior to their use in the split window equation (Gerace et al., 2020). An example comparing Landsat 8 split window surface temperature products generated with and without the low pass filter is shown in Fig. 9. Using output pixel scale that more closely represents the native GSD of the Landsat 8 and 9 TIRS instruments can further reduce aliasing artifacts. Alternative spatial resolutions for Landsat Level 2 surface temperature and emissivity products are discussed in Section 5.

Despite the maturity of the split window algorithm, further investigation is needed to characterize the algorithm-induced discontinuities relative to the single channel approach for long term trend analyses. Initial intercomparisons based on a common dataset indicate that the mean biases of the two retrieval approaches differ by ~ 1.2 K, with the single channel exhibiting warmer surface temperatures on average. These differences are quantified in Section 5.2 and are discussed in the broader context of Landsat continuous long-term TIR record in Section 6.

4.4. Processing enhancements

The C3 Level 2 surface temperature and emissivity processing in C3 will be decoupled from VSWIR-derived Level 2 product generation. This product generation design change removes dependencies on VSWIR atmospheric correction assumptions and auxiliary inputs and enables surface temperature processing for low or no solar illumination conditions (i.e., nighttime imaging). For scenes where Level 2 surface reflectance processing is not applied due to the scene center SZA

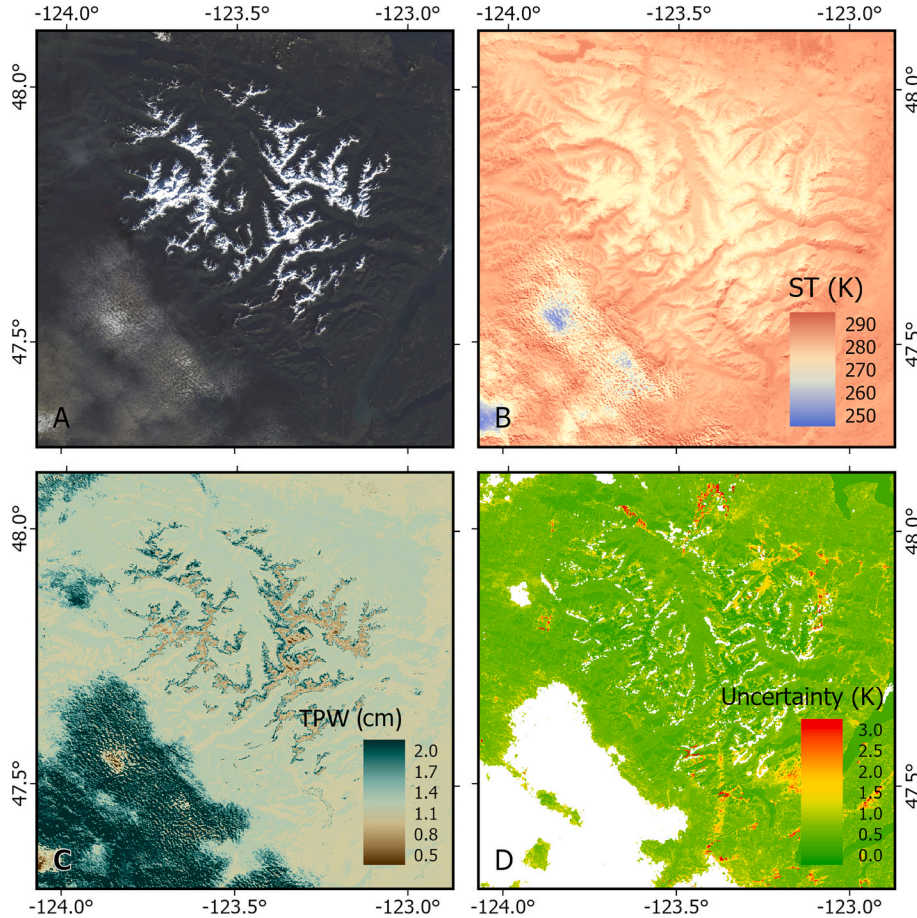


Fig. 8. Example Landsat 9 products showing true color surface reflectance (panel A), split window retrieved surface temperature (panel B), total precipitable water (TPW) (panel C), and split window surface temperature uncertainty derived by incorporating the TPW parameter in a standard error propagation model (panel D). The intermediate TPW layer shows elevated column water content where clouds are as expected. While cloud-contaminated pixels are masked in the uncertainty layer (white), the thin edges of clouds are not always masked revealing the subtle impacts of TPW with slightly elevated uncertainty compared to surrounding pixels. The Landsat 9 image was acquired over the Olympic Peninsula, WA on November 8, 2025, over World Reference System-2 (WRS-2) path 47/row 27.

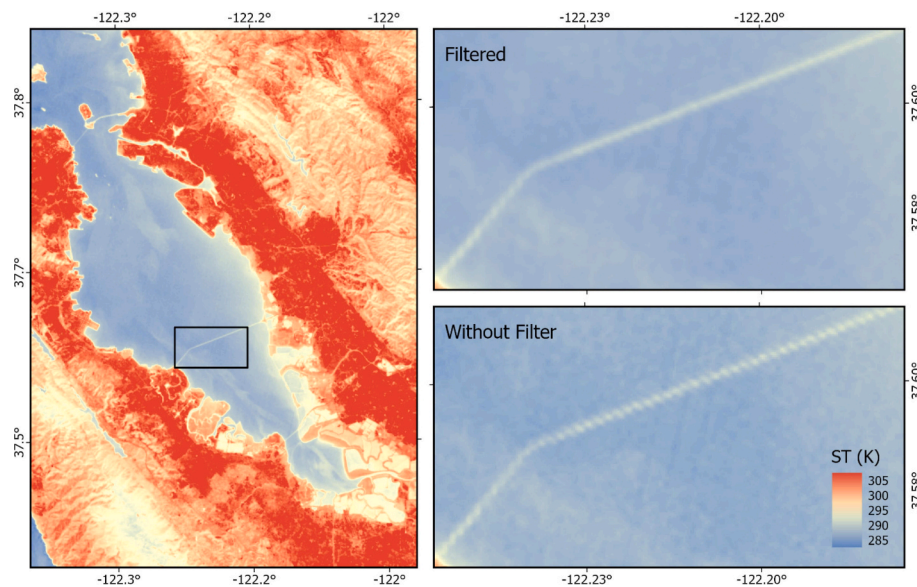


Fig. 9. Example showing the effects of a low pass filter over the Dumbarton Bridge in San Francisco, CA. This Landsat 9 image was acquired on March 25, 2025, over World Reference System-2 (WRS-2) path 44/row 34.

exceeding 76° , NDVI or NDSI values derived from the Level 1 TOA reflectance will support phenology adjustment of ASTER GEDv3 emissivity. Landsat pixel locations lacking ASTER GEDv3 coverage (e.g., Antarctica), scenes with less than 75% ASTER GEDv3 coverage, and nighttime image acquisitions will utilize the monthly CAMEL V003 emissivity that does not require phenology adjustment (refer to Section 4.1.1). This enhancement is particularly important for polar science and application studies where low solar illumination and extended periods of seasonal darkness otherwise limit VSWIR atmospheric correction (Scambos et al., 2000; Moussavi et al., 2020).

Another C3 Level 2 surface temperature and emissivity product enhancement involves increasing the radiometric range limits for all Landsat TIR instruments. In C2, surface temperature retrievals are capped at 373 K, corresponding to the upper limit of radiometric calibration for Landsat TIR instruments. This limitation constraints the representation of higher temperature targets. For C3, instrument-specific lookup tables will be updated to support surface temperature retrievals of up to 500 K, thereby expanding the capabilities for high temperature conditions. Although legacy Landsat TIR sensors radiometrically saturate at temperatures well below 500 K, the revised lookup tables define an extended retrieval domain for converting valid radiance observations to surface temperature. The Landsat 8–9 TIRS $10.8 \mu\text{m}$ band detectors are expected to remain reliable up to ~ 400 K and to saturate beyond this range (Reuter et al., 2015). Pixels affected by radiometric saturation will be assigned a NoData value in both the surface temperature and uncertainty layers.

Finally, C3 single channel algorithm surface temperature retrieval will adopt the Copernicus DEM (COPDEM) (Fahrland et al., 2020; Franks and Rengarajan, 2023), replacing the composite DEM used in C2 (Franks et al., 2020). Although surface temperature is only minimally sensitive to elevation, COPDEM will be used for determining pixel ground elevation in C3 single channel algorithm processing to maintain consistency across the Level 1 and Level 2 processing chain.

5. Accuracy assessment

The performances of the C2 and provisional C3 Landsat surface temperature retrievals were evaluated using Temperature-based (T-based) and Radiance-based (R-based) validation approaches described by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) (Guillevic et al., 2018).

The T-based approach provides the most direct validation but requires accurate in situ surface temperature measurements over thermally homogeneous sites coincident with the satellite overpass (Guillevic et al., 2018). Therefore, only a limited number of operational sites globally meet the requirements for T-based surface temperature validation. The R-based approach relies on radiative closure simulations within the thermal atmospheric window to estimate the surface temperature by comparing the calculated and observed TOA BT at sites with known surface emissivity and representative atmospheric profiles (Guillevic et al., 2018). An advantage of the R-based approach is that it does not require coincident in situ measurements, enabling application across a broader range of sites. However, the R-based approach requires reliable emissivity information from field or laboratory measurements (Hulley et al., 2021). Validation data used in this manuscript are described in detail in Section 5.1.

5.1. Validation data

The geographic distribution of T-based and R-based validation data sources used in this study is shown in Fig. 10. The T-based validation sources include JPL water sites, the National Oceanic and Atmospheric Administration (NOAA) buoy network, and SURFRAD network over the continental United States and Gobabeb validations stations in Namibia. The R-based sites are a network of pseudo-invariant sites in western United States. Each of these sources are described in further detail below.

- a) **JPL Water:** The primary T-based validation data used in this study are from JPL's automated calibration/validation buoys in Lake Tahoe and the Salton Sea. These water sites are considered “flagship” validation locations because the radiometers are calibrated to the National Institute of Standards and Technology (NIST) traceable standards (Hook et al., 2005). The JPL water sites span a broad range of water temperatures ($4\text{--}35^\circ\text{C}$) under varying seasonal atmospheric humidity conditions, providing a robust testbed for TIR measurement accuracy across diverse environmental and atmospheric conditions (Hook et al., 2019). There are four surface sampling buoy stations in Lake Tahoe (tx1, tx2, tx3, and tx4) and one in the Salton Sea. Each buoy is equipped with calibrated TIR radiometers that continuously measure surface radiance, bulk water temperature, and near-surface atmospheric conditions at a high temporal sampling

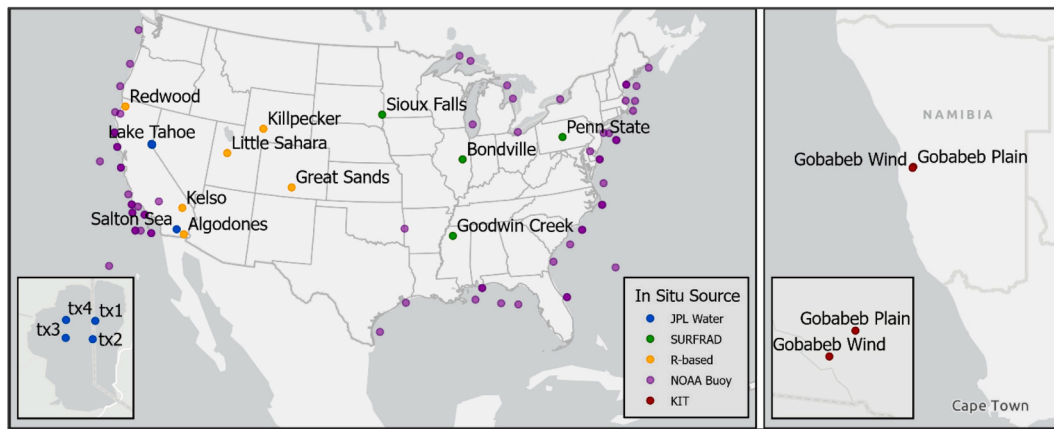


Fig. 10. Geographic distribution of Temperature-based (T-based) and Radiance-based (R-based) validation sites over the continental United States (left) and Namib Desert (right) used for Landsat surface temperature retrieval accuracy assessment. Zoom-in windows show the distribution of Jet Propulsion Laboratory (JPL) water sites in Lake Tahoe in the left panel and Gobabeb stations in the right panel.

rate. The measured radiance includes contributions from the water surface emission and reflected downwelling sky radiance. The sky radiance is estimated using radiative transfer simulations driven by coincident atmospheric profiles and is removed to isolate surface-leaving radiance. Using band effective water emissivity values for Landsat TIR bands, the corrected radiance is converted to surface skin radiometric temperature (Hook et al., 2019).

- b) **NOAA buoy:** NOAA operates a network of coastal, near-shore, and inland buoys across the continental United States through the National Data Buoy Center (NOAA National Data Buoy Center, 2026). Similar to the JPL water sites, these buoy locations provide reliable targets for long-term monitoring of TIR sensor performance because water surfaces exhibit spectrally stable and well-characterized emissivity. For this study, NOAA buoy observations were used to derive reference water surface temperatures from bulk water temperature measurements. Following established approaches, the measured bulk temperature was adjusted to estimate the skin temperature using coincident environmental parameters, including wind speed and atmospheric conditions (Padula, 2008; Eon et al., 2026). Although the NOAA buoys do not provide the same level of radiometric characterization as the JPL water sites, their broad spatial and temporal coverage substantially increases the number of available water validation matchups and complement the dedicated JPL measurements.
- c) **SURFRAD:** The SURFRAD stations are equipped with pyrgeometers and represent a wide range of land surface and atmospheric conditions across the United States (Augustine et al., 2000). The SURFRAD network has been widely used for surface temperature validation because of its continuous measurements and open accessibility (Wang and Liang, 2009). However, previous studies have identified two primary sources of uncertainty. First, pyrgeometers measure hemispherical broadband TIR flux (3.5–50 μm), whereas satellite-derived surface temperature represents a directional quantity, and the shorter-wavelength portion may include reflected solar contamination (Hulley et al., 2021). Second, spatial heterogeneity and shadowing effects have been noted at specific SURFRAD sites, particularly during daytime conditions (Martin et al., 2019; Hulley et al., 2021). Malakar et al. (2018) found that the Sioux Falls, Penn State, Goodwin Creek, and Bondville sites exhibited sufficient spatial homogeneity for Landsat GSD scale. Accordingly, these four SURFRAD sites were selected for this study. At these locations, surface temperature was derived from upwelling and downwelling longwave radiation using the Stefan–Boltzmann law following the procedure described by Malakar et al. (2018), and broadband emissivity values

were obtained from ASTER and MODIS datasets as reported by Guillevic et al. (2018).

- d) **KIT:** The Karlsruhe Institute of Technology (KIT) operates two dedicated surface temperature validation stations located northeast of the Gobabeb Training and Research Centre in the Namib Desert, Namibia. These sites, referred to here as Gobabeb-Wind and Gobabeb-Plain, are situated within a hyperarid environment at the transition between extensive sand dunes and gravel plains (Göttsche and Hulley, 2012). The validation area is characterized by highly homogeneous surfaces composed of gravel, sand, and sparse desiccated vegetation, providing a stable target with unique albedo characteristics. The combination of exceptional spatial homogeneity, temporal stability, and persistently clear-sky conditions makes the Gobabeb sites particularly well suited for long-term calibration and validation of satellite-derived surface temperature products. The primary instrument used for in situ surface temperature measurements at the Gobabeb stations is the KT15.85 IIP precision radiometer (Heitronics GmbH, Wiesbaden, Germany). The radiometers measure thermal infrared radiance between 9.6 and 11.5 μm with a temperature resolution of 0.03 $^{\circ}\text{C}$ and an accuracy of ± 0.3 $^{\circ}\text{C}$ over the relevant temperature range (Theocharous et al., 2019).
- e) **R-based:** The R-based validation sites used in this study include five pseudo-invariant sand dune sites, and one forested field located in the western United States. The pseudo-invariant dune targets have been widely used for calibration and validation of satellite TIR observations because their large, spatially uniform expanses of sand exhibit minimal variability in emissivity (Wan, 2008; Wan and Li, 2008; Hulley et al., 2009; Niclòs et al., 2018; Hulley et al., 2021). In addition, the coniferous redwood forest in northern California provides a vegetated site where persistent canopy offers a relatively stable target for validation. The spectral emissivities of these sites have been characterized by laboratory measurements of sampled materials (Hulley et al., 2009) or by analogue spectral library data for the coniferous forest site (Hulley and Hook, 2012). To estimate the actual surface temperature, site-specific surface emissivity, atmospheric profiles, and satellite-retrieved surface temperature are used as input to the RTTOV forward radiative transfer model to simulate TOA BT in a clear atmospheric window (11–12 μm). The surface temperature is iteratively adjusted until two simulated TOA BT values bound the observed BT. The truth surface temperature is then calculated by interpolating between the simulated and observed BT values (Hulley et al., 2021).

Surface temperature retrievals from Landsat 8 and 9 were evaluated against all five validation sources, whereas Landsat 5 and 7 were

evaluated only using JPL water and SURFRAD sites. The R-based data requires at least two TIR bands which is not applicable to Landsat 5 and 7 observations. Only clear-sky scenes (cloud cover $\leq 20\%$) were included to minimize uncertainties associated with cloud proximity. Selected scenes were further screened for cloud contamination over the validation site using Level 1 pixel QA flags. Following Malakar et al. (2018), matchup statistics were derived from a 3×3 -pixel neighborhood centered on the site location. Finally, outliers were identified and removed on a site-by-site basis using the $1.5 \times$ interquartile range (IQR) criterion applied to the residual distribution (Tukey, 1977). A total of 6182 matchups from Landsat 5–9 were retained for the analysis. Additional site-specific details and number of Landsat matchups per site are provided in Table 4.

5.2. Validation results

Four Landsat surface temperature retrieval configurations were evaluated: (1) the operational C2 surface temperature product which is generated using the single channel algorithm (Earth Resources Observation and Science (EROS) Center, 2020a, 2020b, 2020c) (2) the single channel algorithm based on MODTRAN Version 5.3.2 radiative transfer model with planned C3 surface emissivity improvements, (3) the single channel algorithm based on RTTOV Version 13.1 with the same C3 surface emissivity improvements, and (4) the split window algorithm that also incorporates planned C3 surface emissivity improvements. For brevity, these configurations are referred to as C2 ST, SC MODTRAN, SC RTTOV, and SW, respectively, in validation plots presented in this section. Depending on the Landsat image acquisition date, the operational C2 products use atmospheric reanalysis inputs from MERRA-2, GEOS FP-IT or GEOS-IT (refer to Section 3.3). The provisional C3 configurations (2, 3, and 4) were generated from C2 Level 1 inputs. The single channel implementations were processed using MERRA-2 atmospheric reanalysis to mirror the planned C3 product generation implementation.

Table 4

Summary of surface temperature validation sites used in this study, grouped by validation source and measurement approach. Listed for each site are the elevation above mean sea level, dominant land cover, and number of cloud-free Landsat-in situ matchups.

Source	Site name	Elevation (m)	Land cover	Number of matchups
JPL Water (T-based)	Lake Tahoe	1897	Water	1580
	Salton Sea	−74	Water	367
NOAA (T-based)	NOAA Buoy	0	Water	655
SURFRAD (T-based)	Sioux Falls	473	Grassland	431
	Penn State	376	Cropland	216
	Bondville	230	Grassland	638
	Goodwin Creek	98	Grassland	798
	Gobabeb Plain	466	Gravel plains	196
KIT (T-based)	Gobabeb Wind	435	Gravel plains	222
	Algodones Dunes	94	Quartz sand	329
	Great Sands	2560	Quartz and potassium feldspar sand	150
	Kelso Dunes	800	Quartz, potassium feldspar, and magnetite sand	253
JPL Land (R-based)	Killpecker	2000	Quartz, plagioclase feldspar, magnetite, and epidote sand	146
	Little Sahara	1560	Quartz, plagioclase feldspar, pyroxene, carbonate, and magnetite sand	107
	Redwood Forest	1026	Conifer, redwood trees	94

Fig. 11 shows pairwise scatterplots comparing each Landsat surface temperature retrieval configuration with the corresponding in situ surface temperature matchup quantities for Landsat 8–9 acquisitions. Landsat 5–7 matchups were excluded to enable consistent comparison with the split window configuration, which is only applicable to Landsat 8–9. The Landsat 5–7 results together with Landsat 8–9 results, stratified by validation data source and including error bars representing surface temperature retrieval uncertainty, are provided in Appendix A. For each comparison scenario, the coefficient of determination (R^2), Root Mean Square Error (RMSE), bias, Mean Absolute Error (MAE), and Mean Absolute Percent Error (MAPE) are reported across all matchups.

Relative to the operational C2 surface temperature product (C2 ST), the provisional C3 single channel configuration (SC RTTOV) reduces overall bias by ~ 0.34 K on average, while consistently lowering RMSE, MAE, and MAPE, indicating reductions in both systematic error and overall retrieval error. The magnitude of the error reduction for the SC MODTRAN configuration is comparatively smaller than SC RTTOV because the C2 atmospheric auxiliary data interpolation issues described in Section 3.6.2 remain present in SC MODTRAN configuration. Site-level analyses presented in Appendix A further demonstrates that relative to C2 product, the SC RTTOV reduces bias, RMSE, and MAE across most T-based validation sites. The only exceptions are the Landsat 8–9 NOAA buoy matchups and Landsat 5–7 JPL water matchups, where the magnitude of negative bias increases slightly, accompanied by unchanged or marginally larger MAE values. Comparison against the R-based validation data indicates that the C3 single channel configuration (SC RTTOV) consistently reduces overall bias, RMSE, and MAE relative to C2 product.

Similarly, the provisional C3 split window configuration reduces bias, RMSE, and MAE relative to C2 ST across most T-based validation sites examined, with the exception of NOAA buoy matchups, where bias magnitude, RMSE and MAE are larger than those of the operational C2 product. In contrast, the split window algorithm exhibits a pronounced cold bias against the R-based validations sites, resulting in an overall bias of -0.93 K in Fig. 11. This represents a mean bias shift of ~ 1.55 K relative to C2 surface temperature on average.

Scatterplots in Fig. 11 indicate that both the C2 surface temperature warm bias and the split window cold bias increase with increasing surface temperatures. This temperature-dependent behavior is further illustrated in Fig. 12 for Landsat 8–9 matchups, where bar plots show the mean bias across temperature bins and error bars represent the 95% confidence intervals. While the magnitude of retrieval bias varies with validation dataset, the single channel configurations consistently exhibit a pattern of cold bias at lower temperatures and warm bias at higher temperatures across most sites. Notably, relative to C2 product, the C3 single channel configuration (SC RTTOV) reduces the magnitude of this temperature-dependent bias in most cases, with the most pronounced improvements occurring at higher surface temperatures.

For the split window algorithm, the proposed C3 configuration generally exhibits a persistent cold bias across all temperature ranges, with the exception of SURFRAD matchups. The magnitude of split window and single channel bias divergence is larger at higher temperatures over the R-based and KIT sites, suggesting increased sensitivity of models to emissivity related assumptions at higher temperatures.

Landsat surface temperature retrieval accuracy varies across validation sites, with larger discrepancies primarily attributable to surface and atmospheric environment, sampled temperature ranges, sub-pixel spatial heterogeneity at the Landsat GSD scale (matchups were averaged over a 90×90 -m area), and uncertainties in in situ measurement. Fig. 13 summarizes the bias and RMSE statistics for each retrieval configuration and validation source. The RMSE at the JPL and NOAA buoy sites remained consistently low across all evaluated surface temperature retrieval configurations, demonstrating the high spatial homogeneity of the water targets, well-characterized emissivity, and well-calibrated radiometers. In contrast, RMSE values are generally larger for the SURFRAD and KIT sites due to uncertainties in assumed and/or

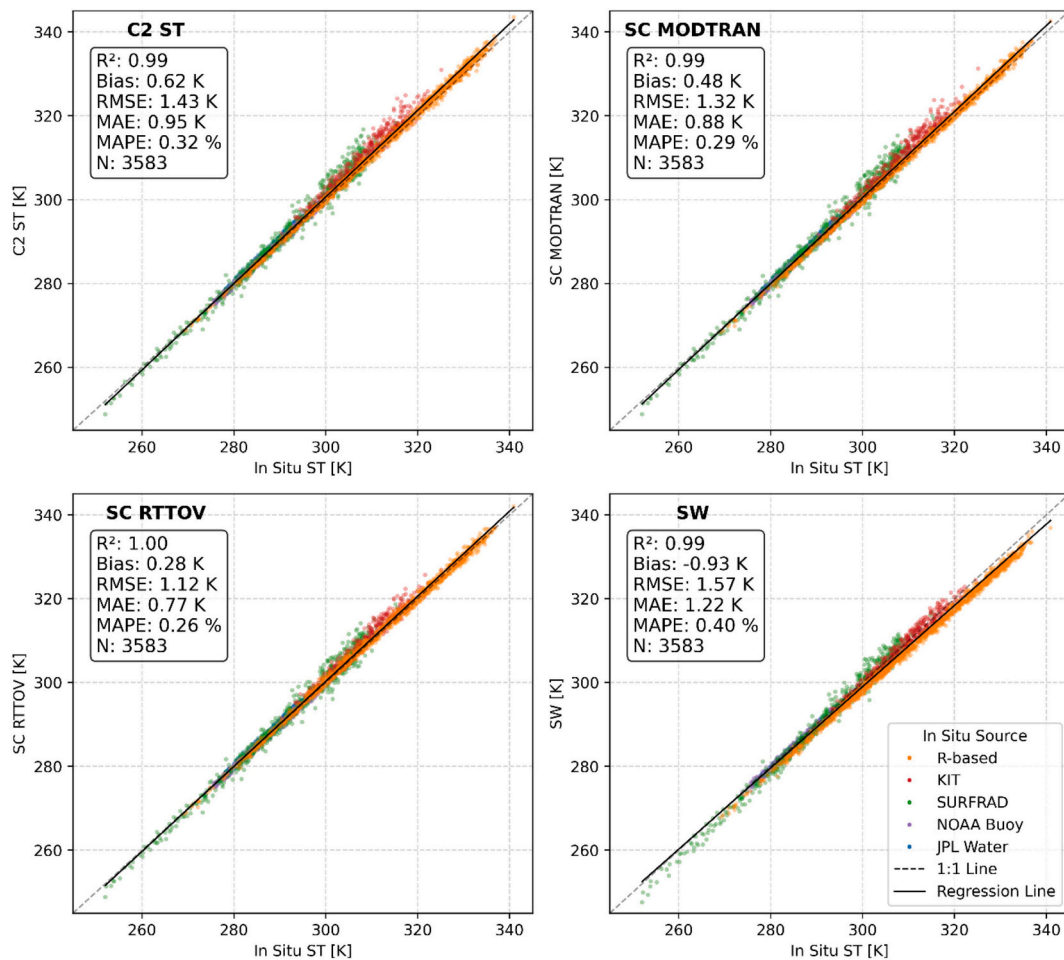


Fig. 11. Scatterplots from 3583 Landsat 8–9 matchups for Collection 2 (C2) single channel product (configuration 1; top left), single channel algorithm using MODerate resolution atmospheric TRANsmission (MODTRAN) with improved Collection 3 (C3) surface emissivity (configuration 2; top right), single channel algorithm using Radiative Transfer for Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) (RTTOV) and improved C3 surface emissivity (configuration 3; bottom left) and split window algorithm incorporating C3 surface emissivity retrieval (configuration 4; bottom right) algorithms compared to the matchup validation data.

measured emissivity, the greater heterogeneity of land surfaces, and seasonal changes in surface conditions. For the R-based validation sites, the single channel configurations exhibit considerably lower RMSE than the split window approach.

Consistent with RMSE results, measured biases are strongly site dependent. The JPL and NOAA water sites demonstrate very small biases across all single channel configurations. The largest bias magnitude over water occurs for Landsat 8 matchups over the JPL water sites and is likely attributable to residual stray light effects. The SURFRAD and KIT sites exhibit a tendency towards warm bias for the single channel retrieval configurations, with largest bias observed at the KIT sites. The R-based validation sites also exhibit a warm bias on average for the single channel surface temperature retrieval configurations. In contrast, the split window configuration exhibits a persistent cold bias across all validation sources, although its magnitude is relatively small over T-based sites. However, when evaluated against the R-based sites, the Landsat 8 and 9 split window configuration exhibits a substantial cold bias exceeding 2 K on average. As mentioned earlier, this systematic discrepancy, coupled with the relatively consistent performance over T-based sites, indicates that the split window formulation is more sensitive to emissivity characterization and inter-band radiometric differences at higher temperature. With the exception of the single channel retrievals at the KIT sites and the split window retrievals evaluated against the R-based sites, both the C3 single channel (SC RTTOV) and split window configurations meet or exceed the ~ 1 K accuracy target over both land

and water surfaces, indicating increased accuracy relative to the operational C2 product for the matchups evaluated in this study.

Fig. 14 further illustrates the distribution of Landsat 8 and 9 retrieval biases using violin plots. The JPL water and NOAA buoy sites show stable performance for Landsat 9, with narrower interquartile ranges (25th–75th percentiles are shown using black dashed lines) that are largely contained within ± 1 K accuracy bounds (red dashed lines), whereas Landsat 8 exhibits a broader spread in bias. The wider distributions for Landsat 8 biases are consistent with known differences in the radiometric performance of the Landsat 8 and 9 TIRS instruments. Bias distributions over the SURFRAD and KIT sites display larger dispersion, indicating the combined effects of spatial heterogeneity at the Landsat spatial scale and uncertainties associated with the reference measurements, including pyrometer-derived surface temperature in SURFRAD sites. Despite the increased spread, the C3 single channel configuration consistently shifts the median bias closer to zero relative to the operational C2 product. Validation against R-based sites shows similar distribution shapes for Landsat 8 and 9, characterized by a persistent negative bias in the split window configuration as noted earlier.

5.3. Algorithmic continuity

Assessing the consistency of surface temperature retrievals across different instrument generations and algorithmic approaches is essential for preserving continuity within the Landsat TIR data record and

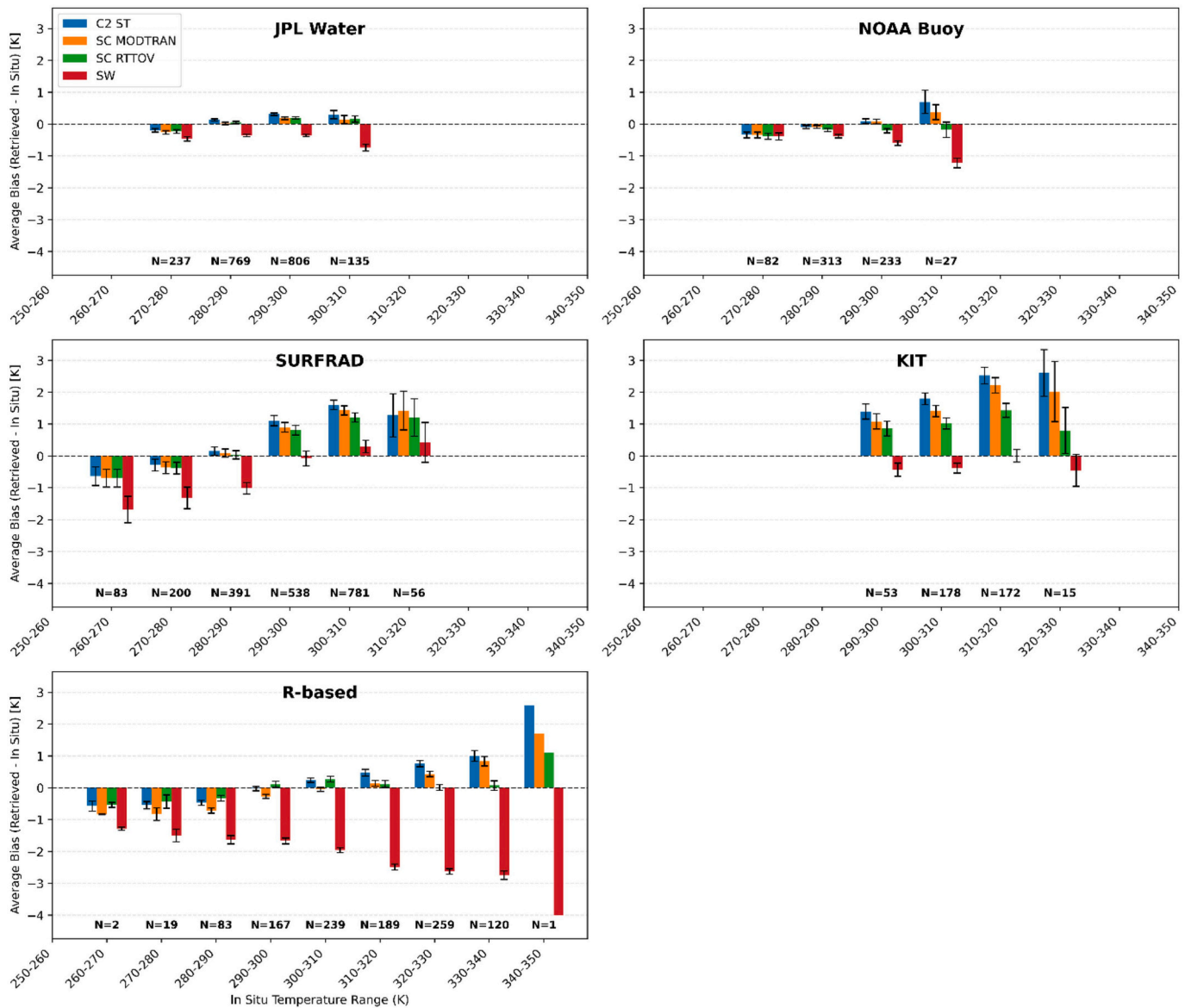


Fig. 12. Average bias and associated 95% confidence intervals for Landsat 8–9 surface temperature retrievals, binned in 10 K intervals. The bar plots highlight temperature-dependent behavior in the retrieval errors, with single channel configurations exhibiting a cold bias at temperatures below ~ 290 K and a warm bias at higher temperatures. In contrast, the split window approach shows a persistent cold bias, with magnitude varying depending on validation data source.

enabling robust analysis of long-term climate and environmental trends. To evaluate algorithmic continuity, annual surface temperature timeseries were generated by applying both the single channel and split window algorithms to identical Landsat 8–9 Level 1 inputs. Fig. 15 illustrates an example over the JPL water sites which highlights the implications of using different algorithms for long-term surface temperature retrieval stability. Differences between the split window and single channel retrievals are more pronounced for Landsat 8, whereas Landsat 9 demonstrates increased stability consistent with reduced stray light sensitivity. For both instruments, algorithm differences increase with higher temperature conditions, with the largest deviations occurring during summer periods characterized by elevated atmospheric water vapor and ambient temperatures. This temperature-dependent divergence is consistent with the increased sensitivity of Landsat 8 TIRS measurements to residual stray light effects and atmospheric correction uncertainties under high TIR radiance conditions (Montanaro et al., 2014b). Across the water sites, single channel algorithm retrievals are systematically warmer than split window results, with mean residuals of 0.86 K for Landsat 8 and 0.09 K for Landsat 9. As

noted in Section 5.2, this algorithmic bias is often larger over the land-based SURFRAD, KIT, and R-based validation sites.

Rigorous uncertainty characterization and quantification techniques are critical for interpreting algorithm-dependent differences across the Landsat surface temperature data record. The current split window uncertainty estimates are generally larger at higher temperatures than those from the single channel algorithm (Figs. 15–20). These temperature-dependent uncertainty patterns demonstrate the importance of incorporating uncertainty information when analyzing long-term trends or combining retrievals derived from different algorithms. Further investigation is needed to better understand the sources of increased split window uncertainty at higher temperatures, and to harmonize uncertainty estimation approaches between algorithms. In particular, extending the TPW-based uncertainty framework to the single channel algorithm would likely increase consistency in uncertainty retrieval and support smoother integration of surface temperature products across Landsat missions and processing collections.

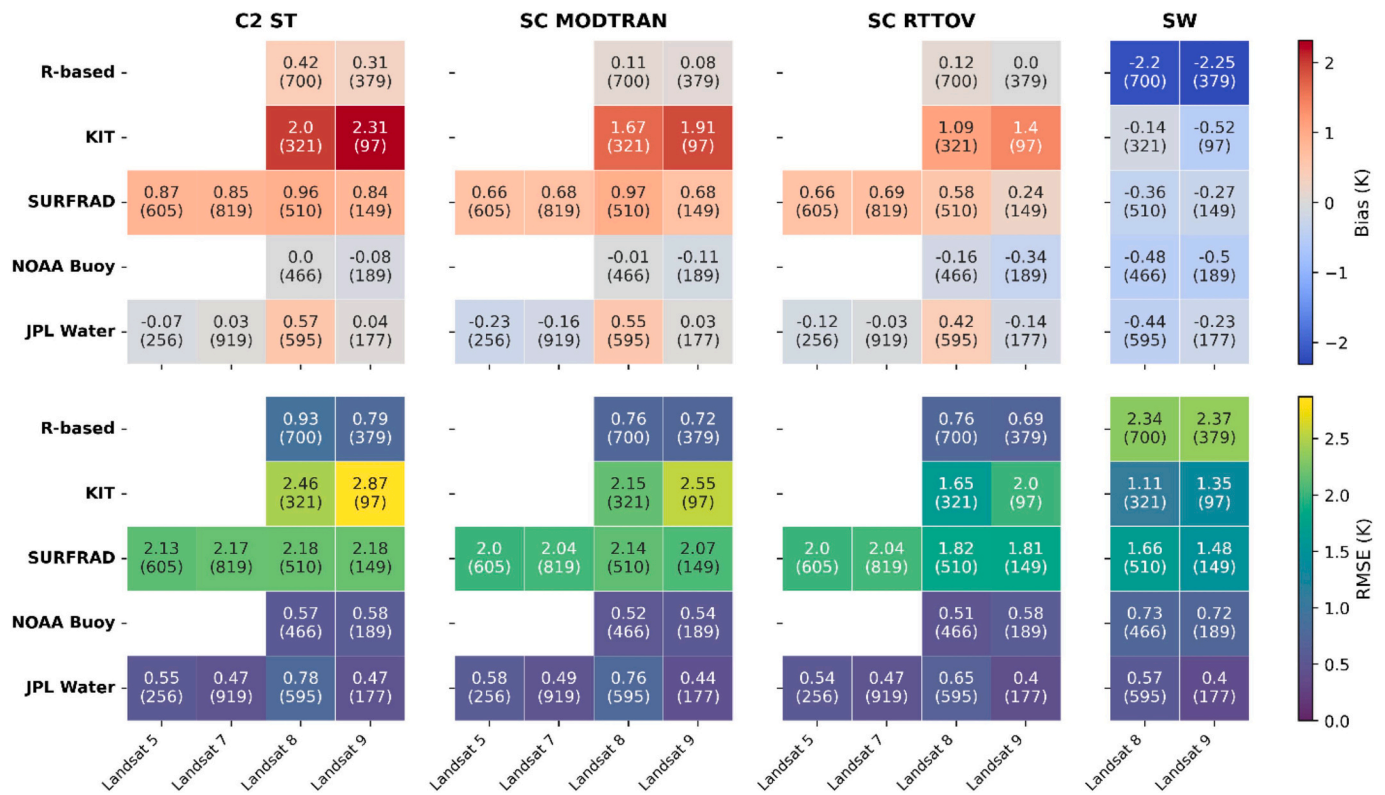


Fig. 13. Summary of surface temperature bias (top row) and Root Mean Square Error (RMSE) (bottom row) statistics by Landsat mission and validation data source. The number of matchups in each scenario is provided in parentheses.

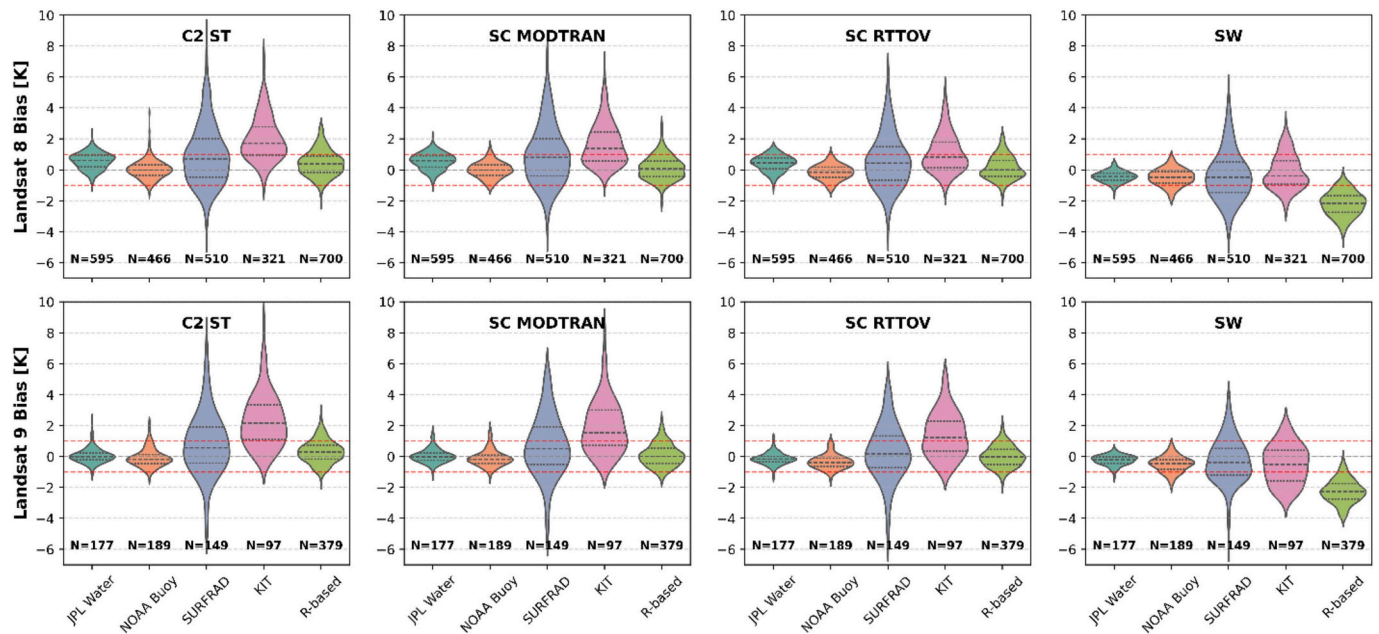


Fig. 14. Violin plot showing the distribution of Landsat 8 and 9 surface temperature bias for each retrieval configuration. Dashed red lines indicate the ± 1 K bias threshold. Dashed black lines within each violin denote the first quartile, median, and third quartile of the bias distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6. Discussion and future direction

The discussion below frames the proposed C3 enhancements in terms of key characteristics of Earth observation data products, including accuracy, continuity, scalability, and usability, consistent with the product-oriented perspective advocated by [Radeloff et al. \(2024\)](#). This

framing highlights the algorithmic advances introduced for Landsat surface temperature and emissivity retrieval and their implications for the operational maturity, interoperability, and long-term value of the Landsat thermal infrared record.

From an accuracy perspective, despite advancements in implementation of split window algorithm for Landsat 8–9 surface

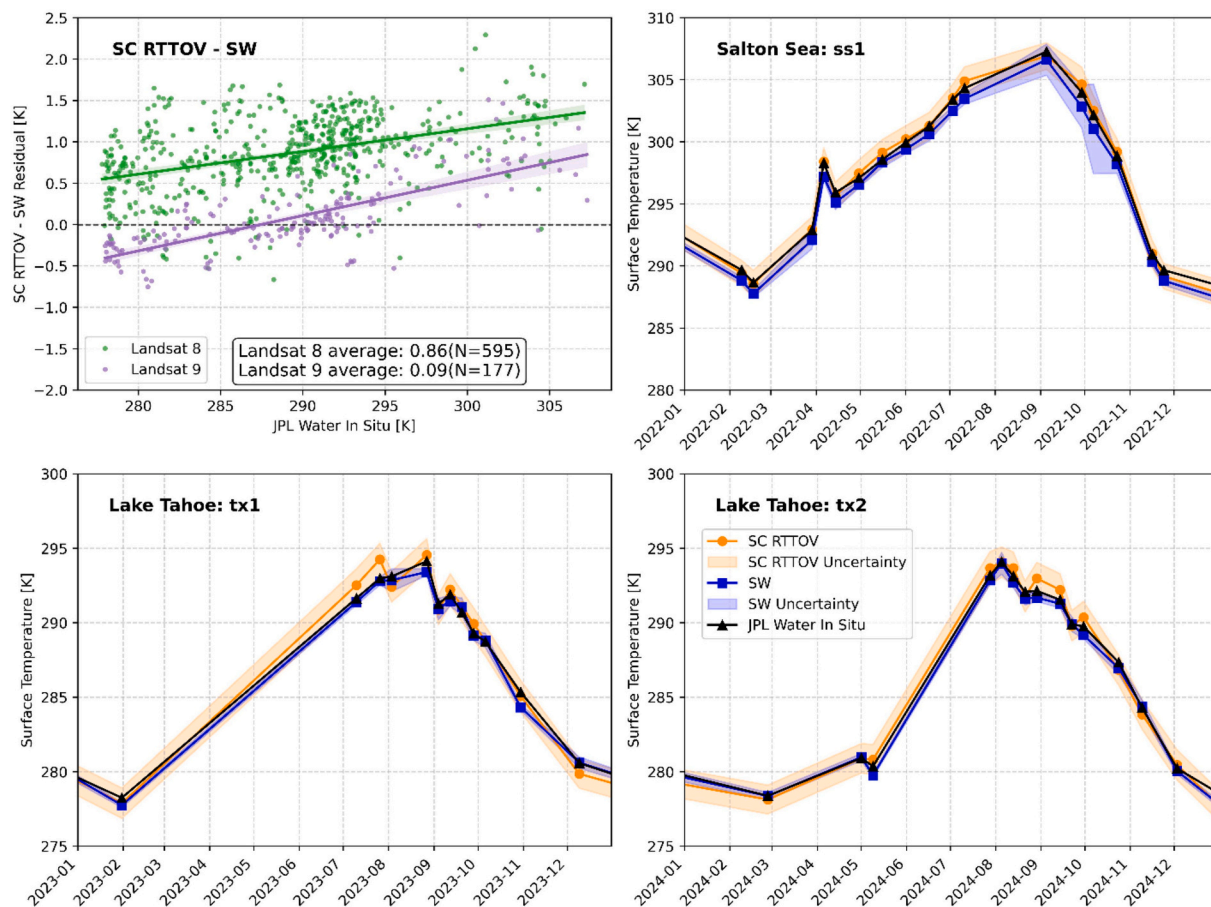


Fig. 15. Systematic differences between the single channel and split window surface temperature retrievals (top left) show warmer retrievals for the single channel algorithm relative to split window across all four JPL water sites. Regression lines and 95% confidence intervals for data from each mission are displayed using a consistent color scheme. Corresponding annual surface temperature and uncertainty timeseries for the Salton Sea (top right), Lake Tahoe (tx1) (bottom left), and Lake Tahoe (tx2) (bottom right) Jet Propulsion Laboratory (JPL) water sites during 2022–2024 illustrate the same bias over an annual time period.

temperature retrieval and associated uncertainty characterization, operational adoption of this approach for C3 remains under evaluation. Introducing the split window algorithm for Landsat 8–9 while continuing to rely on the single channel approach for Landsat 4–7 has the potential to introduce discontinuities in the long-term Landsat Level 2 surface temperature data record due to differences in retrieval physics and sensitivity to environmental conditions. Timeseries analyses over water targets indicate that algorithmic differences up to ~ 1.5 K may occur depending on Landsat mission, surface heating conditions, and atmospheric humidity, even under consistent surface emissivity assumptions and radiative transfer modeling. In addition, validation over land sites suggests that relative to single channel, split window retrievals may exhibit an average cold bias exceeding 2 K under extreme temperature conditions. However, these differences should be interpreted in the context of the distinct assumptions underlying each validation strategy, as discussed in Section 5.

Different applications may benefit more from one or the other approach. For example, the split window algorithm generally provides improved atmospheric water vapor compensation under humid conditions, whereas the single channel method may offer greater consistency with heritage Landsat missions TIR data record. One potential direction for preserving long-term continuity is development of adaptive atmospheric correction strategies, such as algorithm selection or blending based on land cover type, emissivity variability, or prevailing climate regime. Further investigation of hybrid retrieval frameworks supported by robust uncertainty characterization and cross-validation across

independent validation sites could help balance absolute accuracy increases with the need for temporal consistency across the Landsat Level 2 surface temperature data record.

The R-based datasets provide an established and widely adopted approach for validating satellite-derived surface temperature products, particularly over high-temperature, pseudo-invariant desert surfaces where direct in situ measurements are impractical (Guillevic et al., 2018; Hulley and Hook, 2012). Unlike T-based validation, however, the reference temperatures are not measured directly but are derived from dual-channel TIR observations combined with assumed stable surface emissivity and external atmospheric state information. Some degree of methodological dependence exists between the R-based reference temperatures and the single channel retrieval, which may partially explain the closer agreement between the single channel retrieval and the R-based reference temperatures observed in this study. Regardless of the absolute performance of the individual retrieval configurations, achieving consistency between the single channel and split window products remains an important objective for preserving the continuity of the long-term Landsat Level 2 surface temperature record in C3. A similar divergence between the two approaches at high temperatures is also evident at the KIT Gobabeb sites; however, in this case, the split window retrievals exhibit closer agreement with the reference temperatures. These findings indicate that additional investigation is warranted to better understand and reconcile the differences between the single channel and split window algorithms, particularly over extreme temperature conditions.

The lower accuracy of the single channel using MODTRAN configuration relative to RTTOV in Section 5.2 is likely attributable to the atmospheric auxiliary data interpolation issues described in Section 3.6.2, which remain present in the MODTRAN implementation. Consequently, the observed differences should not be interpreted as evidence of inferior radiative transfer performance by MODTRAN. In principle, MODTRAN is expected to provide accurate simulations because it uses a physics-based line-by-line treatment of gaseous absorption, whereas RTTOV relies on a computationally efficient parameterized formulation trained on representative atmospheric conditions. Radiative transfer intercomparison exercises have generally demonstrated excellent agreement between modern thermal infrared radiative transfer models, with line-by-line approaches often being the reference standard against which fast models such as RTTOV are evaluated (Saunders et al., 2007). Therefore, the differences reported here are partially driven by implementation-specific auxiliary data handling rather than fundamental limitations of the MODTRAN framework.

An important consideration for the long-term Landsat surface temperature record is the consistency of retrievals across successive sensor generations. Maintaining continuity in the multi-decadal Landsat thermal record is particularly important because even small algorithmic differences can influence trend analyses and change detection applications. Although the analyses presented in this manuscript primarily assess algorithm performance relative to independent reference datasets, cross-sensor continuity also depends on the radiometric stability of the underlying thermal observations. Quarterly assessments conducted by the USGS EROS Calibration and Validation Center of Excellence (ECCOE) indicate that the thermal infrared instruments onboard Landsat 4–9 have maintained stable radiometric performance relative to their onboard blackbody calibrators throughout their operational lifetimes (e.g., Haque et al., 2025). Nevertheless, subtle differences in sensor spectral response functions, noise characteristics, and instrument configurations may still influence surface temperature consistency. Future work could therefore leverage same-day acquisitions during Landsat mission overlap periods to directly quantify cross-sensor agreement and further characterize the continuity of the multi-decadal Landsat surface temperature record.

Beyond accuracy and continuity, scalability and operational sustainability are critical considerations for future Landsat thermal products. With a planned launch readiness date in the early 2030s, Landsat 10 will continue the TIR observation record established by heritage Landsat missions. Depending on the outcome of mission architecture deliberations, new TIR bands centered at 8.3, 8.6, and 9.1 μm would enable direct retrieval of land surface spectral emissivity using a multi-channel TIR algorithm, such as TES (Gillespie et al., 1998; Hulley et al., 2021), or Optimal Estimation (OE) algorithm that leverages a-priori knowledge of the surface and atmospheric parameters (Kassar et al., 2021). For surfaces with well-characterized spectral emissivity, such as open water, the generalized split window algorithm remains relevant due to reduced sensitivity to emissivity retrieval noise (Qin et al., 2001). Atmospheric correction of Landsat 10 TIR observations may consider an ensemble of algorithms tailored to specific surface type. This would balance the use of any new TIR spectral bands while preserving continuity across heritage Landsat missions for multi-decadal Earth surface studies. Landsat C3 STE product interoperability with other medium resolution TIR missions that are planned to operate in the late 2020s and 2030s, including LSTM, TRISHNA, and EAGLE-TIR, will be enabled through international partnerships and technical interchanges to promote the collaborative coordination recommended by Turpie et al. (2023). Landsat 10 TIR observations, combined with data from these complementary missions, could also contribute to a new global emissivity dataset, providing a critical reference for Landsat 8 and 9 STE product generation into the 2030s. In particular, the spectral surface emissivity from Landsat 10, in combination with coincident VSWIR observations, will support medium-resolution geological mapping and mineral characterization of newly exposed rocky surfaces in high-

latitude regions of Greenland and Antarctica. Extended Landsat acquisitions in polar regions further enhance opportunities to monitor Earth's surface change under low solar illumination conditions (Scambos et al., 2024). Additionally, nighttime surface temperature retrievals further enable applications such as UHI monitoring (Arellano and Roca, 2021) and volcanological assessments (Nádudvari et al., 2020). Decoupling surface temperature processing from surface reflectance also allows retrievals from Landsat 8 and 9 TIRS-only observations that lack coincident VSWIR data.

Product usability is also influenced by how information is packaged, documented, and delivered to users. The C3 Landsat Level 2 STE product will include surface temperature, surface temperature uncertainty, and spectral surface emissivity raster layers that correspond to the number of TIR spectral bands for each Landsat mission. For Landsat 4–9, an emissivity source layer will document whether external sources or prescribed quantities were used to determine Landsat surface emissivity. For Landsat 10, this layer may be reserved to report the atmospheric correction algorithms applied for each surface type. The single channel algorithm's intermediate atmospheric parameters available in C2 that includes spectral atmospheric transmittance, upwelling and downwelling radiance, and TIR radiance, will be removed. These layers were included in the Landsat provisional C1 surface temperature product and continued in C2 to support processing traceability on an operational basis. Now that C2 surface temperature products are available as a pre-processed global inventory, and C3 research and development has increased algorithm maturity, the removal of these intermediate atmospheric characterization layers will improve consistency across Landsat missions and reduce data storage needs. In addition, the distance-to-cloud layer available in the C2 product will be removed because it will no longer be used for uncertainty characterization in C3.

Although Landsat TIR instruments acquire data with a GSD ranging from 60 to 120 m over the data record, C1 and C2 Landsat 4–9 Level 1 calibrated TIR radiance and Level 2 surface temperature products have been generated with a 30-m pixel size. This is achieved by resampling raw TIR image data to the VSWIR grid using cubic convolution and Akima interpolation (USGS, 2025). This enabled users to leverage both VSWIR and TIR data in their analyses without additional resampling (Landsat Update, Volume 4, Issue 1, 2010). However, in addition to the aliasing artifacts due to detector layout discussed in Section 3.1, the combination of resampling and the spatial performance of the instruments introduce other challenges in the representation of high spatial frequency information, such as narrow features and sharp edges (Moik, 1980). The substantial upsampling of Landsat TIR data may lead users to underestimate these effects because the spatial scale over which signal is integrated at contributing instrument elements is much coarser than the pixel scale of USGS-provided Level 1 and Level 2 TIR data products (Wenny et al., 2015; Eftekharzadeh Kay et al., 2024). Studies that have a specific focus on the thermal characteristics of ecological boundaries, urban microclimates, or narrow waterways would be at the greatest risk of impacts (e.g., Martí-Cardona et al., 2019; Reek et al., 2025; Alonzo et al., 2025).

For C3, the USGS is considering adopting upper-left pixel referencing to improve geometric alignment with other missions such as Sentinel-2 and simplify pixel nesting for spatial resolutions that are multiples of 30 m, reducing the need for the legacy 30-m TIR resampling. Using the native TIR GSDs of 120 m for Landsat 4–5 TM, 60 m for Landsat 7 ETM+ and the planned Landsat 10, and 100 m for Landsat 8–9 TIRS introduces challenges for pixel alignment and multidecadal time series analysis. To address these challenges, the USGS is evaluating a common 60-m spatial resolution that is native to Landsat 7 and the Landsat 10 design and requires pixel value replication for Landsat 4–5 TM and Landsat 8–9 TIRS. For Landsat 8–9 TIRS, this approach yields an effective native spatial resolution of 120 m (rather than 90 m), which more closely aligns with instrument's nominal 100-m GSD and broader 200-m Point Spread Function (PSF) width, thereby reducing oversampling relative to sensor's effective spatial response (Wenny et al., 2015; Holmes et al., 2024;

Eon et al., 2024).

In addition to changes in pixel referencing and TIRS spatial resolution discussed above, C3 is expected to incorporate several Level 1 and Level 2 product improvements with implications for Level 2 surface temperature and emissivity retrievals. While TIR spectral bands will continue to be calibrated using coastal and inland buoy measurements (Barsi et al., 2014), C3 will cross-calibrate VSWIR reflectance data from all Landsat missions to corresponding Landsat 9 Operational Land Imager (OLI) spectral bands as the absolute radiometric reference. Although the sensitivity of Level 2 surface temperature and emissivity products to these changes has yet to be quantified, planned improvements in Level 2 surface reflectance processing combined with use of Landsat 9 as the absolute radiometric reference for the complete Landsat data record are expected to increase consistency in surface reflectance derived spectral indices used for temporal emissivity adjustment. Updates to QA are also anticipated to improve the detection of snow/ice and water presence on a per-pixel basis and will likely aid in more reliable prescribed emissivity values.

In Section 3.3, multiple external atmospheric profile sources that are used for C2 Landsat 4–9 surface temperature product generation were described. With potential transition to a generalized split window algorithm for C3 Landsat 8 and 9 TIR processing and the decommissioning of Landsat 7 in 2025, latency of atmospheric auxiliary data may no longer be a determining factor for the single channel algorithm. Therefore, a single external source of atmospheric auxiliary data source can be adopted for C3 to improve temporal processing consistency. Both MERRA-2 and the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) span the full temporal range of Landsat 4–7 TIR observations and are suitable for Landsat. NASA's upcoming MERRA-3 with updated GEOS atmospheric components will also be evaluated upon availability. The USGS will assess available sources prior to C3 reprocessing to identify the optimal atmospheric auxiliary dataset for Landsat data record.

The C2 Landsat Level 2 surface temperature product currently meets the Threshold-level requirements of the Committee of Earth Observation Satellites (CEOS) ARD Surface Temperature Product Family Specification (PFS) version 5.0 (Lewis et al., 2018; Siqueira et al., 2019). With the C3 TIR processing enhancements described in this paper, which also include improved metadata machine readability, the Landsat Level 2 STE product is expected to meet the Target-level CEOS PFS specifications for Surface Temperature ARD. While no specific CEOS PFS currently exists for spectral surface emissivity, emissivity correction is explicitly called for within the CEOS Surface Temperature PFS which is currently satisfied in the C2 ST product and will continue to be satisfied by the Landsat C3 STE product.

Prior to the official release of C3, the USGS plans to offer limited early access to a beta version of the Level 2 STE product by request only to members of the Landsat Science Team and interested stakeholders within the TIR remote sensing community. The primary goal of early access to C3 STE beta products is to increase transparency, solicit structured user feedback, and address any potential issues before initiating a large-scale data record reprocessing campaign. The user feedback will inform C3 development and implementation procedures without extending the public release schedule, and will ensure a robust, well-characterized Level 2 STE product to enable a broad range of scientific and operational applications in the 2030s.

7. Conclusions

This paper has summarized the planned enhancements and advancements for the Landsat C3 Level 2 STE product being developed by the USGS in collaboration with NASA JPL and RIT. Building on lessons learned from the global provision of surface temperature products in C2, Landsat C3 will represent a substantial evolution in surface emissivity characterization, atmospheric correction techniques, uncertainty quantification, and the end-to-end processing architecture of the Landsat 4–9

TIR data record. These advancements also establish a clear on-ramp for Landsat 10 TIR observations and data products.

Accuracy assessments against a diverse set of validation datasets demonstrate measurable improvements in retrieval accuracy and consistency for the proposed C3 atmospheric correction configurations relative to operational C2 product, while also highlighting performance can vary with environmental conditions and land cover type. These results demonstrate the importance of continued evaluation of surface temperature retrieval accuracy and estimated uncertainty across a broad range of climate and land cover conditions across different biome and climate regions, particularly during warm season periods and extreme atmospheric conditions; these results also reinforce the value of large, well-characterized validation datasets with calibrated instruments. In closing, the algorithmic refinements and processing enhancements will position the USGS in-development Landsat C3 Level 2 STE products on a path to operational maturity in support of a broad range of scientific, operational, and climatological applications on multi-decadal timescales.

CRediT authorship contribution statement

Saeed Arab: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Chase Mueller:** Writing – review & editing, Visualization, Validation, Software, Investigation, Formal analysis. **Christopher J. Crawford:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Danika Wellington:** Writing – review & editing, Investigation. **Ray Dittmeier:** Writing – review & editing, Software. **Glynn Hulley:** Writing – review & editing, Software, Methodology, Data curation, Conceptualization. **Aaron Gerace:** Writing – review & editing, Software. **Calli Jenkerson:** Writing – review & editing, Supervision. **Simon Hook:** Writing – review & editing, Methodology, Data curation. **Amirhossein Hassanzadeh:** Writing – review & editing, Software. **Rehman Eon:** Writing – review & editing, Software. **Cody Anderson:** Writing – review & editing, Supervision, Project administration. **Pat Scaramuzza:** Writing – review & editing. **Frank M. Götsche:** Data curation, Writing – review & editing.

Funding

This research was primarily supported by the USGS Core Science Systems National Land Imaging Program through the Land Satellite Data Systems Research and Development project (GX26ED00EDK1A00). KBR work was conducted through USGS EROS Technical Support Services Contract (Contract No. 140G0121D001). A portion of this work was carried out at the JPL, California Institute of Technology, under a USGS-NASA Interagency Agreement (No. 4820001730). RIT work was supported by the USGS under a Cooperative Agreement (No. G24AC00644).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Mac Friedrichs (KBR, Contractor to the USGS EROS Center) for his valuable feedback. The authors also thank Matthew Fagan (University of Maryland, Baltimore County) for insights into current use cases of Landsat thermal data that may be sensitive to the spatial performance of the Landsat TIR instruments and the resampling algorithm applied to the data. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U. S. Government.

Appendix A. Detailed validation scatter plots

Fig. 16 through Fig. 22 show the scatterplots of retrieved and in situ surface temperatures for validation sources used in this study. These figures are included to provide site-specific surface temperature retrieval accuracy information for Fig. 11.

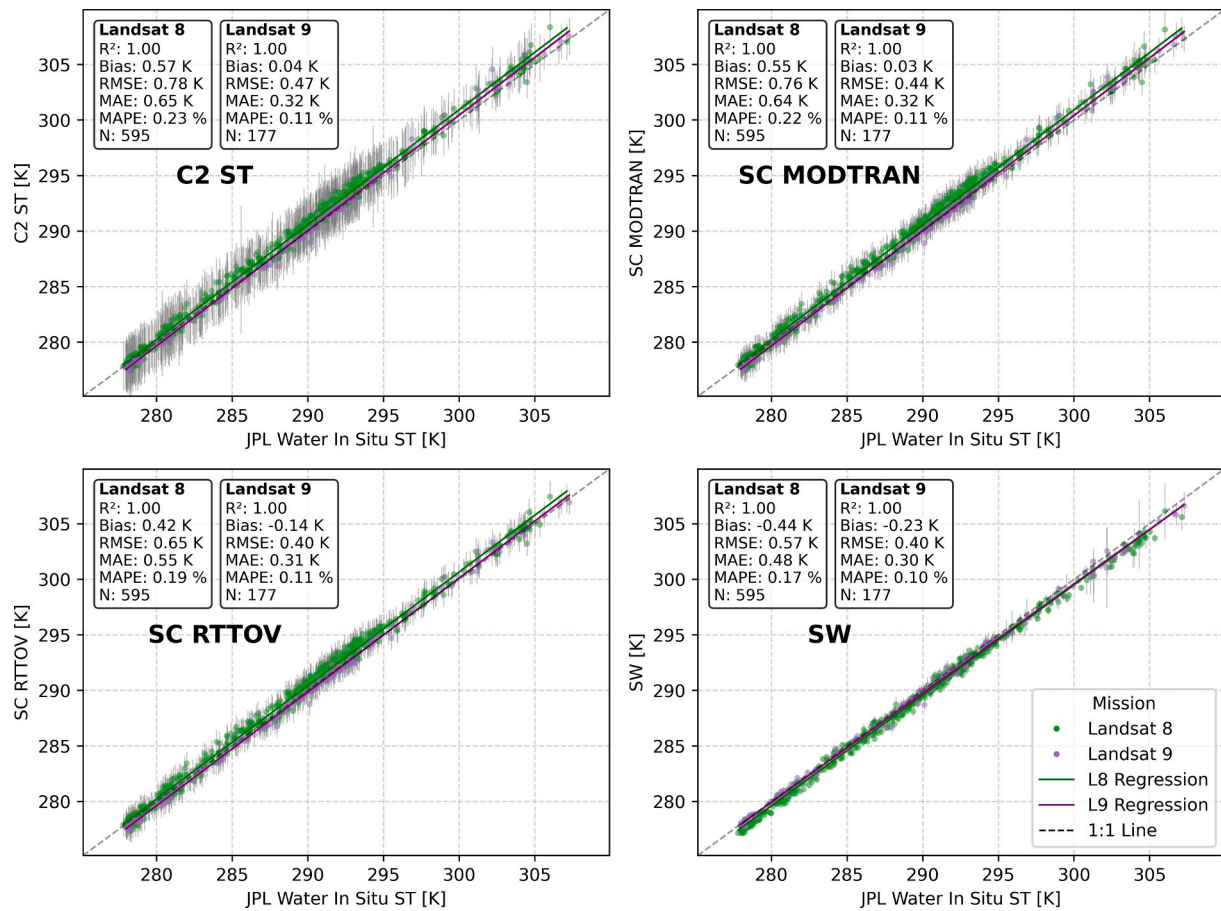


Fig. 16. Scatterplots comparing Landsat 8–9 surface temperature retrievals against in situ measurements from Jet Propulsion Laboratory (JPL) water validation sites.

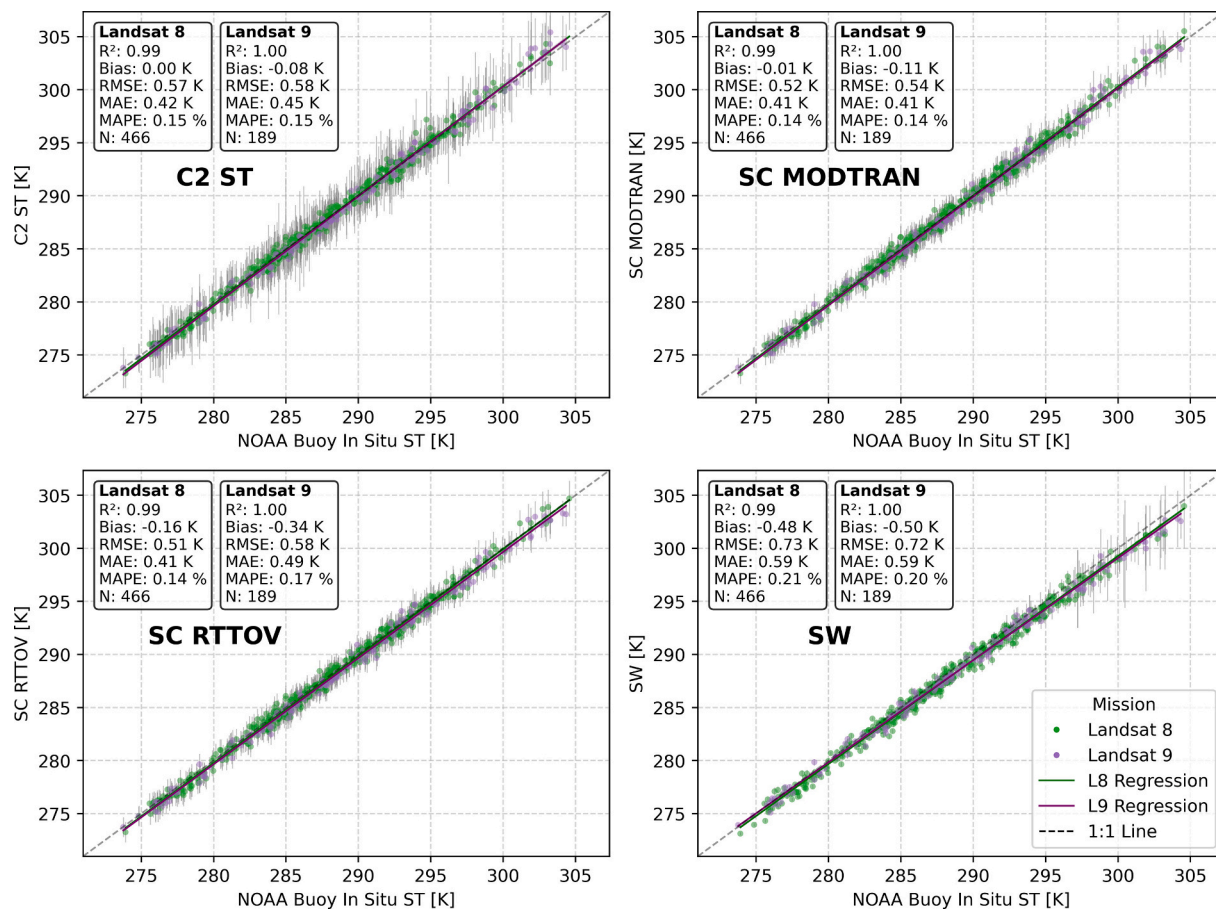


Fig. 17. Scatterplots comparing Landsat 8–9 surface temperature retrievals against in situ measurements from the National Oceanic and Atmospheric Administration (NOAA) buoy validation sites.

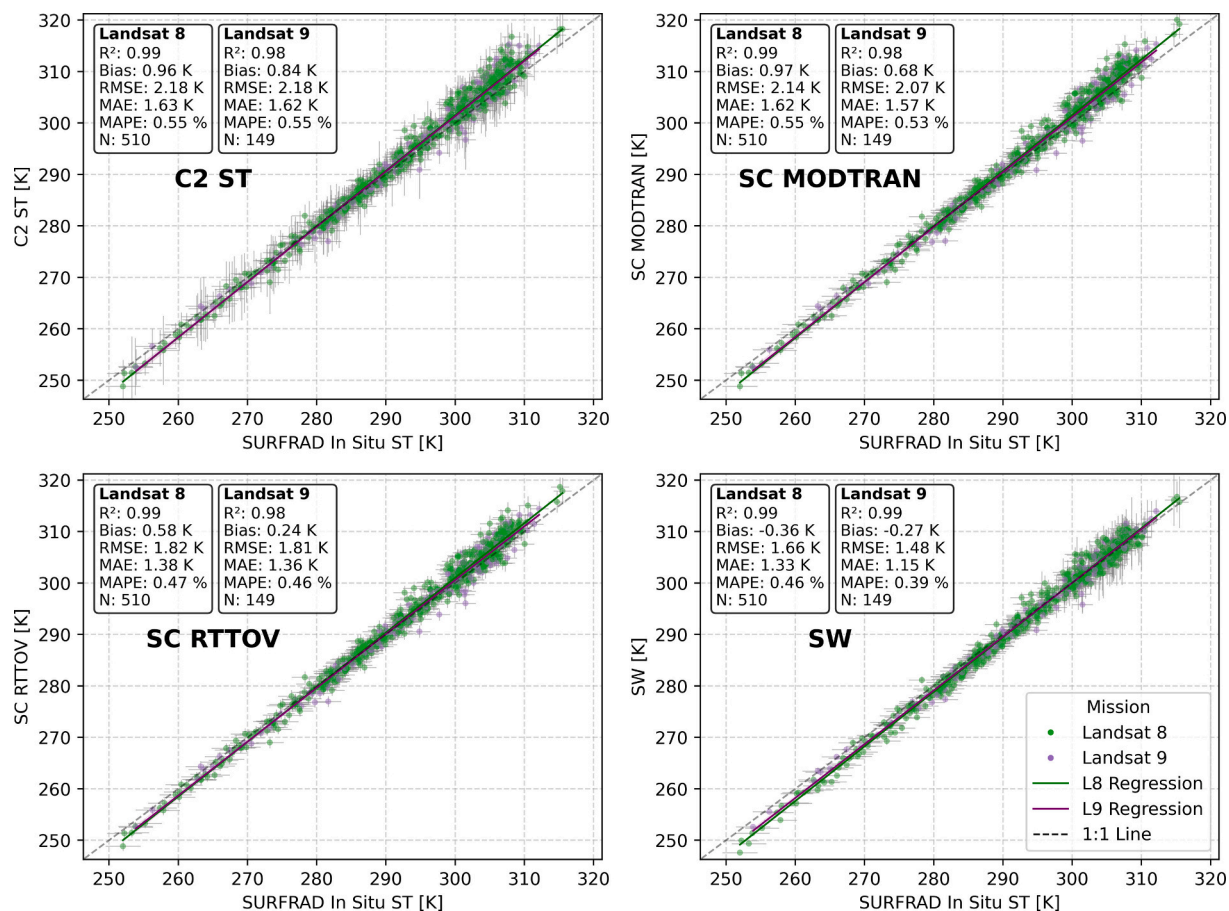


Fig. 18. Scatterplots comparing Landsat 8–9 surface temperature retrievals against in situ measurements from National Oceanic and Atmospheric Administration (NOAA) Surface Radiation Budget (SURFRAD) sites.

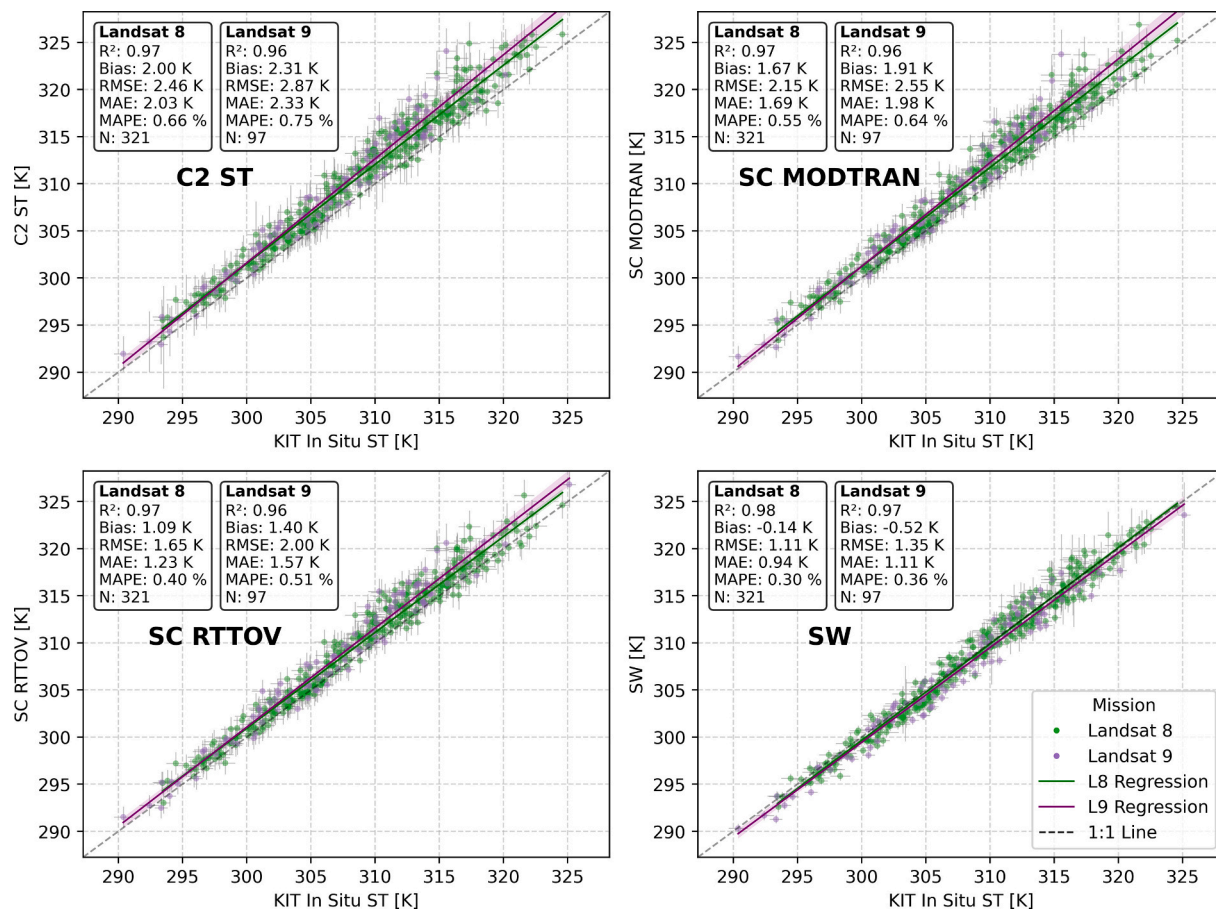


Fig. 19. Scatterplots comparing Landsat 8–9 surface temperature retrievals against in situ measurements from Karlsruhe Institute of Technology (KIT) validation sites.

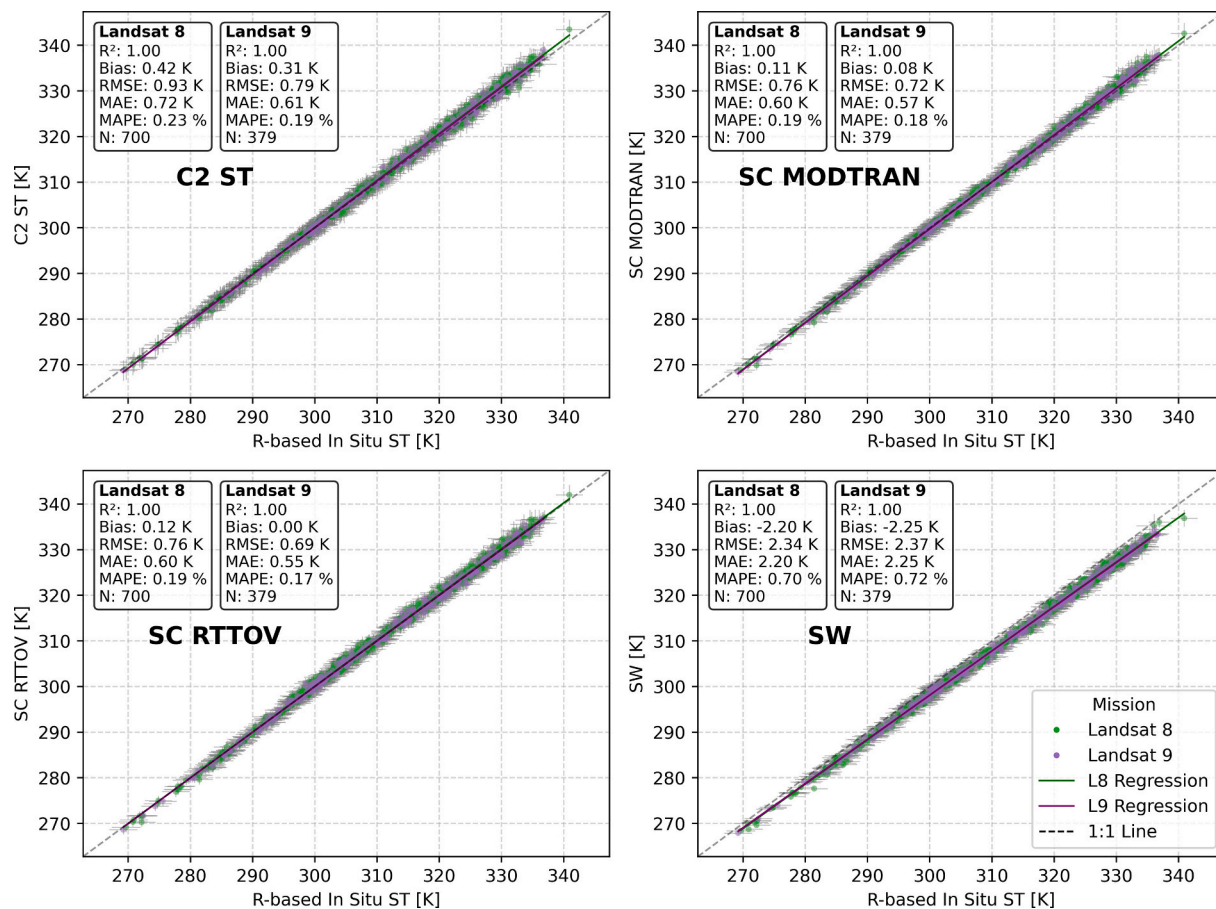


Fig. 20. Scatterplots comparing Landsat 8–9 surface temperature retrievals against Radiance-based (R-based) validation data.

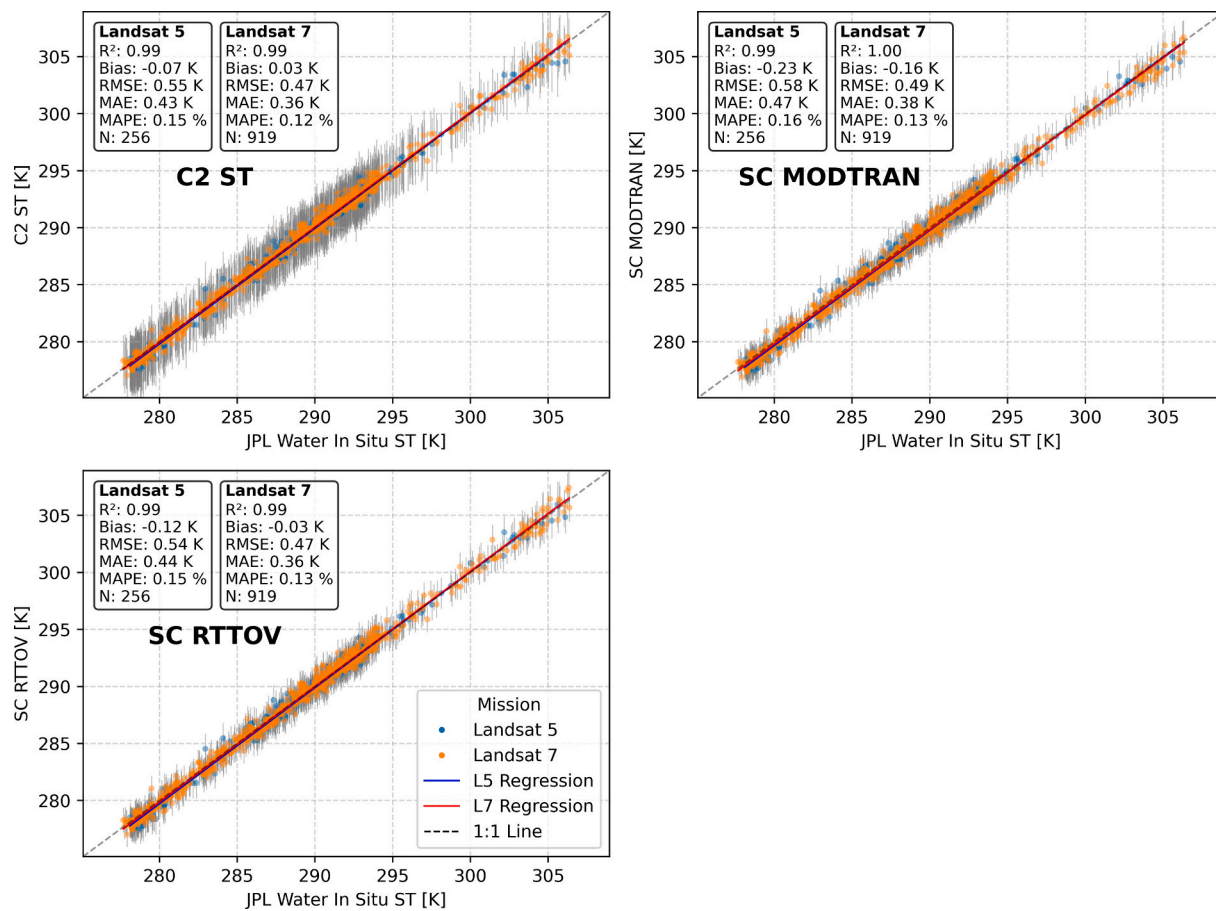


Fig. 21. Scatterplots comparing Landsat 5–7 surface temperature retrievals against in situ measurements from Jet Propulsion Laboratory (JPL) water validation sites.

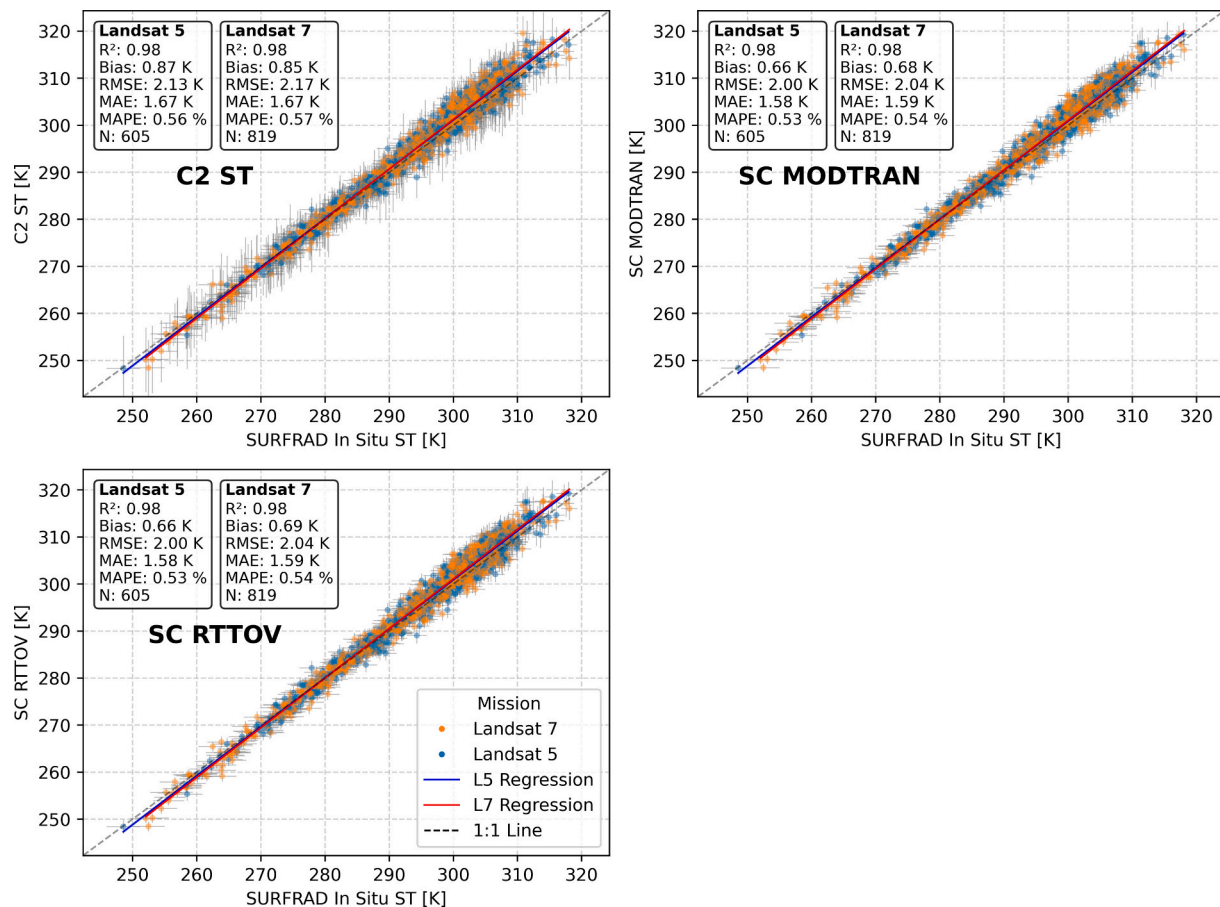


Fig. 22. Scatterplots comparing Landsat 5–7 surface temperature retrievals against in situ measurements from National Oceanic and Atmospheric Administration (NOAA) Surface Radiation Budget (SURFRAD) sites.

Data availability

Data will be made available on request.

References

- Abrams, M., Hook, S., Ramachandran, B., 2002. ASTER User Handbook, Version 2, 4800. Jet Propulsion Laboratory, p. 135. https://lpdaac.usgs.gov/documents/262/ASTER_User_Handbook_v2.pdf (accessed December 10, 2025).
- Allison, R.S., Johnston, J.M., Craig, G., Jennings, S., 2016. Airborne optical and thermal remote sensing for wildfire detection and monitoring. *Sensors* 16 (8), 1310. <https://doi.org/10.3390/s16081310>.
- Alonzo, M., Ibsen, P.C., Locke, D.H., 2025. Urban trees and cooling: a review of the recent literature (2018 to 2024). *Arboric. Urban For.* 51. <https://doi.org/10.48044/jauf.2025.023>.
- Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sens. Environ.* 122, 50–65. <https://doi.org/10.1016/j.rse.2011.08.025>.
- Arellano, B., Roca, J., 2021. Remote sensing and night time urban heat island. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* 43, 15–22. <https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-15-2021>.
- Arvidson, T., Barsi, J., Jhabvala, M., Reuter, D., 2013. Landsat and thermal infrared imaging. In: *Thermal Infrared Remote Sensing: Sensors, Methods, Applications*. Springer Netherlands, Dordrecht, pp. 177–196. https://doi.org/10.1007/978-94-007-6639-6_9.
- Augustine, J.A., DeLuisi, J.J., Long, C.N., 2000. SURFRAD—a national surface radiation budget network for atmospheric research. *Bull. Am. Meteorol. Soc.* 81 (10), 2341–2358. [https://doi.org/10.1175/1520-0477\(2000\)081%3C2341:SANSRB%3E2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081%3C2341:SANSRB%3E2.3.CO;2).
- Barsi, J.A., Barker, J.L., Schott, J.R., 2003. An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument. In: *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, 5. IEEE, pp. 3014–3016. <https://doi.org/10.1109/IGARSS.2003.1294665>.
- Barsi, J.A., Schott, J.R., Hook, S.J., Raqueno, N.G., Markham, B.L., Radocinski, R.G., 2014. Landsat-8 thermal infrared sensor (TIRS) vicarious radiometric calibration. *Remote Sens.* 6 (11), 11607–11626. <https://doi.org/10.3390/rs6111607>.
- Barsi, J.A., Markham, B.L., Montanaro, M., Gerace, A., Hook, S.J., Schott, J.R., Morfitt, R., 2017. Landsat-8 TIRS thermal radiometric calibration status. In: *Earth Observing Systems XXII*, 10402. SPIE, pp. 461–471. <https://doi.org/10.1117/12.2276045>.
- Barsi, J.A., Montanaro, M., Thome, K., Raqueno, N.G., Hook, S., Anderson, C.H., Micijevic, E., 2022. Early radiometric performance of landsat-9 thermal infrared sensor. In: *Earth Observing Systems XXVII*, 12232. SPIE, pp. 223–233. <https://doi.org/10.1117/12.2634058>.
- Becker, F., Li, Z.L., 1990. Towards a local split window method over land surfaces. *Remote Sens.* 11 (3), 369–393. <https://doi.org/10.1080/01431169008955028>.
- Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Shettle, E.P., 2006, May. MODTRAN5: 2006 update. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII*, 6233. SPIE, pp. 508–515. <https://doi.org/10.1117/12.665077>.
- Bernard, F., Manolis, I., Barat, I., Alamañac, A.B., Taboada, M.S., Mingorance, P., Holgueras, D.G., 2023, October. The Copernicus land surface temperature monitoring (LSTM) mission: design, technology and status. In: *Sensors, Systems, and Next-Generation Satellites XXVII*, 12729. SPIE, pp. 41–55. <https://doi.org/10.1117/12.2679705>.
- Borbas, E.E., Hulley, G., Feltz, M., Knuteson, R., Hook, S., 2018. The combined ASTER MODIS emissivity over land (CAMEL) part 1: methodology and high spectral resolution application. *Remote Sens.* 10 (4), 643. <https://doi.org/10.3390/rs10040643>.
- Borbas, E., Hook, S., Hulley, G., Loveless, M., Knuteson, R., 2023. Combined ASTER and MODIS Emissivity database over Land (CAMEL) Emissivity Monthly Global 0.05Deg V003 [Dataset]. NASA Land Processes Distributed Active Archive Center. <https://doi.org/10.5067/MEASURES/LSTE/CAM5K30EM.003> (accessed 31 December 2025).
- Cawse-Nicholson, K., Townsend, P.A., Schimel, D., Assiri, A.M., Blake, P.L., Buongiorno, M.F., SBG Algorithms Working Group, 2021. NASA's surface biology and geology designated observable: a perspective on surface imaging algorithms. *Remote Sens. Environ.* 257, 112349. <https://doi.org/10.1016/j.rse.2021.112349>.

- Cawse-Nicholson, K., Luvall, J.C., Hook, S., Lee, C. (Eds.), 2026. High Spatio-Temporal Spectral Thermal Remote Sensing: Research and Applications. CRC Press. <https://doi.org/10.1201/9781003458562>.
- Cook, M., Schott, J.R., Mandel, J., Raqueno, N., 2014. Development of an operational calibration methodology for the Landsat thermal data archive and initial testing of the atmospheric compensation component of a land surface temperature (LST) product from the archive. *Remote Sens.* 6 (11), 11244–11266. <https://doi.org/10.3390/rs6111244>.
- Cooper, B.L., Salisbury, J.W., Killen, R.M., Potter, A.E., 2002. Midinfrared spectral features of rocks and their powders. *J. Geophys. Res. Planets* 107 (E4), 1. <https://doi.org/10.1029/2000JE001462>.
- Crawford, C.J., Roy, D.P., Arab, S., Barnes, C., Vermote, E., Hulley, G., Zahn, S., 2023. The 50-year Landsat collection 2 archive. *Sci. Remote Sens.* 8, 100103. <https://doi.org/10.1016/j.srs.2023.100103>.
- Dwyer, J.L., Roy, D.P., Sauer, B., Jenkinson, C.B., Zhang, H.K., Lymburner, L., 2018. Analysis ready data: enabling analysis of the Landsat archive. *Remote Sens.* 10 (9), 1363. <https://doi.org/10.3390/rs10091363>.
- Dyba, K., Ermida, S., Ptak, M., Piekarczyk, J., Sojka, M., 2022. Evaluation of methods for estimating lake surface water temperature using Landsat 8. *Remote Sens.* 14 (15), 3839. <https://doi.org/10.3390/rs14153839>.
- Earth Resources Observation and Science (EROS) Center, 2020a. Landsat 8-9 Operational Land Imager / Thermal Infrared Sensor Level-2, Collection 2 [Dataset]. U.S. Geological Survey. <https://doi.org/10.5066/P9OGBGM6>.
- Earth Resources Observation and Science (EROS) Center, 2020b. Landsat 7 Enhanced Thematic Mapper Plus Level-2, Collection 2 [Dataset]. U.S. Geological Survey. <https://doi.org/10.5066/P9C7113B>.
- Earth Resources Observation and Science (EROS) Center, 2020c. Landsat 4-5 Thematic Mapper Level-2, Collection 2 [Dataset]. U.S. Geological Survey. <https://doi.org/10.5066/P9IAXOVV>.
- Eftekhazadeh Kay, S., Wenny, B.N., Thome, K.J., Yarahmadi, M., Lampkin, D.J., Tahersima, M.H., Voskani, N., 2024. Assessment of spatial characterization metrics for on-orbit performance of Landsat 8 and 9 thermal infrared sensors. *Remote Sens.* 16 (19), 3588. <https://doi.org/10.3390/rs16193588>.
- Eon, R., Wenny, B.N., Poole, E., Eftekhazadeh Kay, S., Montanaro, M., Gerace, A., Thome, K.J., 2024. Landsat 9 thermal infrared sensor-2 (TIRS-2) pre- and post-launch spatial response performance. *Remote Sens.* 16 (6), 1065. <https://doi.org/10.3390/rs16061065>.
- Eon, R.S., Gerace, A., Montanaro, M., 2026. A forward modeling approach for the vicarious validation of the thermal infrared sensor on-board Landsat 8 and 9. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* <https://doi.org/10.1109/JSTARS.2026.3670640>.
- Fahrland, E., Jacob, P., Schrader, H., Kahabka, H., 2020. Copernicus Digital Elevation Model Product Handbook. Airbus Defence and Space—Intelligence, Potsdam, Germany, 2024-06. https://dataspace.copernicus.eu/sites/default/files/media/files/2024-06/geo1988-copernicusdem-spe-002_producthandbook_i5.0.pdf (accessed 20 December 2025).
- Feltz, M., Borbas, E., Knuteson, R., Hulley, G., Hook, S., 2018. The combined ASTER and MODIS emissivity over land (CAMEL) global broadband infrared emissivity product. *Remote Sens.* 10 (7), 1027. <https://doi.org/10.3390/rs10071027>.
- Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley Jr., R.D., Beckmann, T., Laue, B., 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* 194, 379–390. <https://doi.org/10.1016/j.rse.2017.03.026>.
- Franks, S., Rengarajan, R., 2023. Evaluation of Copernicus DEM and comparison to the DEM used for Landsat collection-2 processing. *Remote Sens.* 15 (10), 2509. <https://doi.org/10.3390/rs15102509>.
- Franks, S., Storey, J., Rengarajan, R., 2020. The new landsat collection-2 digital elevation model. *Remote Sens.* 12 (23), 3909. <https://doi.org/10.3390/rs12233909>.
- Galve, J.M., Sánchez, J.M., García-Santos, V., González-Piqueras, J., Calera, A., Villodre, J., 2022. Assessment of land surface temperature estimates from Landsat 8-TIRS in a high-contrast semiarid agroecosystem. Algorithms intercomparison. *Remote Sens.* 14 (8), 1843. <https://doi.org/10.3390/rs14081843>.
- Gelaro, R., McCarty, W., Suárez, M.J., Todling, M., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* 30 (14), 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Gerace, A., Montanaro, M., 2017. Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard Landsat 8. *Remote Sens. Environ.* 191, 246–257. <https://doi.org/10.1016/j.rse.2017.01.029>.
- Gerace, A., Kleynhans, T., Eon, R., Montanaro, M., 2020. Towards an operational, split window-derived surface temperature product for the thermal infrared sensors onboard Landsat 8 and 9. *Remote Sens.* 12 (2), 224. <https://doi.org/10.3390/rs12020224>.
- Gillespie, A., Rokugawa, S., Matsunaga, T., Cothorn, J.S., Hook, S., Kahle, A.B., 1998. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. *IEEE Trans. Geosci. Remote Sens.* 36 (4), 1113–1126. <https://doi.org/10.1109/36.700995>.
- Gök, D.T., Scherler, D., Wulf, H., 2024. Land surface temperature trends derived from Landsat imagery in the Swiss Alps. *Cryosphere* 18 (11), 5259–5276. <https://doi.org/10.5194/tc-18-5259-2024>.
- Göttsche, F.M., Hulley, G.C., 2012. Validation of six satellite-retrieved land surface emissivity products over two land cover types in a hyper-arid region. *Remote Sens. Environ.* 124, 149–158.
- Guillevis, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Camacho, F., 2018. Land surface temperature product validation best practice protocol. Version 1.1. In: *Best Practice for Satellite-Derived Land Product Validation*, p. 60. <https://doi.org/10.5067/doc/ceoswgcv/lpv/1st.001>.
- Guillevis, P.C., Olioso, A., Hook, S.J., Fisher, J.B., Lagouarde, J.P., Vermote, E.F., 2019. Impact of the revisit of thermal infrared remote sensing observations on evapotranspiration uncertainty—a sensitivity study using AmeriFlux data. *Remote Sens.* 11 (5), 573. <https://doi.org/10.3390/rs11050573>.
- Haque, M.O., Hasan, M.N., Shrestha, A., Rengarajan, R., Lubke, M., Steinwand, D., Pinto, C.T., 2025. ECCOE Landsat Quarterly Calibration and Validation Report—Quarter 4, 2024 (No. 2025-1036). US Geological Survey. <https://doi.org/10.3133/ofr20251036>.
- Hassanzadeh, A., Mancini, R., Gerace, A., Eon, R., Montanaro, M., 2025. An Uncertainty Workflow for Landsat TIRS Split Window Surface Temperature Products Incorporating Atmospheric Water Vapor. arXiv preprint. <https://doi.org/10.48550/arXiv.2511.12729>.
- Herrick, C., Steele, B.G., Brentup, J.A., Cottingham, K.L., Ducey, M.J., Lutz, D.A., Palace, M.W., Thompson, M.C., Trout-Haney, J.V., Weathers, K.C., 2023. lakeCoSTR: a tool to facilitate use of Landsat collection 2 to estimate lake surface water temperatures. *Ecosphere* 14 (1), e4357. <https://doi.org/10.1002/ecs2.4357>.
- Holmes, T.R., Poulter, B., McCorkel, J., Jennings, D.E., Wu, D.L., Efremova, B., Hook, S.J., 2024. On-orbit spatial performance characterization for thermal infrared imagers of Landsat 7, 8, and 9, ECOSTRESS and CTI. *J. Geophys. Res. Biogeosci.* 129 (2), e2023JG007506. <https://doi.org/10.1029/2023JG007506>.
- Hook, S., Hulley, G., 2019. ECOSTRESS Land Surface Temperature and Emissivity Daily L2 Global 70 m V001 [Dataset]. NASA Land Processes Distributed Active Archive Center. <https://doi.org/10.5067/ECOSTRESS/ECO2LSTE.001> (accessed 02 January 2026).
- Hook, S.J., Gabell, A.R., Green, A.A., Kealy, P.S., 1992. A comparison of techniques for extracting emissivity information from thermal infrared data for geologic studies. *Remote Sens. Environ.* 42, 123–135. [https://doi.org/10.1016/0034-4257\(92\)90096-3](https://doi.org/10.1016/0034-4257(92)90096-3).
- Hook, S.J., Clodius, W.B., Balick, L., Alley, R.E., Abtahi, A., Richards, R.C., Schladow, S.G., 2005. In-flight validation of mid-and thermal infrared data from the multispectral thermal imager (MTI) using an automated high-altitude validation site at Lake Tahoe CA/NV, USA. *IEEE Trans. Geosci. Remote Sens.* 43 (9), 1991–1999. <https://doi.org/10.1109/TGRS.2005.853191>.
- Hook, S.J., Cawse-Nicholson, K., Barsi, J., Radocinski, R., Hulley, G.C., Johnson, W.R., Markham, B., 2019. In-flight validation of the ECOSTRESS, Landsats 7 and 8 thermal infrared spectral channels using the Lake Tahoe CA/NV and Salton Sea CA automated validation sites. *IEEE Trans. Geosci. Remote Sens.* 58 (2), 1294–1302. <https://doi.org/10.1109/TGRS.2019.2945701>.
- Hulley, G.C., Hook, S.J., 2010. Generating consistent land surface temperature and emissivity products between ASTER and MODIS data for earth science research. *IEEE Trans. Geosci. Remote Sens.* 49 (4), 1304–1315. <https://doi.org/10.1109/TGRS.2010.2063034>.
- Hulley, G.C., Hook, S.J., 2012. A radiance-based method for estimating uncertainties in the atmospheric infrared sounder (AIRS) land surface temperature product. *J. Geophys. Res. Atmos.* 117 (D20). <https://doi.org/10.1029/2012JD018102>.
- Hulley, G.C., Hook, S.J., Baldrige, A.M., 2009. Validation of the North American ASTER land surface emissivity database (NAALSED) version 2.0 using pseudo-invariant sand dune sites. *Remote Sens. Environ.* 113 (10), 2224–2233. <https://doi.org/10.1016/j.rse.2009.06.005>.
- Hulley, G.C., Hook, S.J., Abbott, E., Malakar, N., Islam, T., Abrams, M., 2015. The ASTER global emissivity dataset (ASTER GED): mapping Earth's emissivity at 100 meter spatial scale. *Geophys. Res. Lett.* 42 (19), 7966–7976. <https://doi.org/10.1002/2015GL065564>.
- Hulley, G., Hook, S., Fisher, J., Lee, C., 2017, July. ECOSTRESS, a NASA Earth-ventures instrument for studying links between the water cycle and plant health over the diurnal cycle. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, pp. 5494–5496. <https://doi.org/10.1109/IGARSS.2017.8128248>.
- Hulley, G.C., Göttsche, F.M., Rivera, G., Hook, S.J., Freepartner, R.J., Martin, M.A., Johnson, W.R., 2021. Validation and quality assessment of the ECOSTRESS level-2 land surface temperature and emissivity product. *IEEE Trans. Geosci. Remote Sens.* 60, 1–23. <https://doi.org/10.1109/TGRS.2021.3079879>.
- Hulley, G., Dehaene, Q., Agatep, A., Shreevastava, A., Runkel, M., 2026. Commentary: mapping and mitigating urban extreme heat using NASA's high-resolution thermal infrared data. *J. Urban Aff.* 1–9. <https://doi.org/10.1080/07352166.2025.2607382>.
- Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: the Landsat data continuity mission. *Remote Sens. Environ.* 122, 11–21. <https://doi.org/10.1016/j.rse.2011.08.026>.
- ISO, I., & OIML, B, 1993. *Guide to the Expression of Uncertainty in Measurement*. Aenor, Madrid, Spain.
- Jedlovec, G.J., 1990. Precipitable water estimation from high-resolution split window radiance measurements. *J. Appl. Meteorol.* (1988-2005) 863–877. [https://doi.org/10.1175/1520-0450\(1990\)029%3C0863:PWEFHR%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1990)029%3C0863:PWEFHR%3E2.0.CO;2).
- Jiménez-Muñoz, J.C., Sobrino, J.A., 2003. A generalized single-channel method for retrieving land surface temperature from remote sensing data. *J. Geophys. Res. Atmos.* 108 (D22). <https://doi.org/10.1029/2003JD003480>.
- Jiménez-Muñoz, J.C., Cristóbal, J., Sobrino, J.A., Soria, G., Ninyerola, M., Pons, X., 2008. Revision of the single-channel algorithm for land surface temperature retrieval from Landsat thermal-infrared data. *IEEE Trans. Geosci. Remote Sens.* 47 (1), 339–349. <https://doi.org/10.1109/TGRS.2008.2007125>.
- Jimenez-Munoz, J.C., Sobrino, J.A., Skoković, D., Mattar, C., Cristóbal, J., 2014. Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data.

- IEEE Geosci. Remote Sens. Lett. 11 (10), 1840–1843. <https://doi.org/10.1109/LGRS.2014.2312032>.
- Kassar, J.E., Carbajal Henken, C., Preusker, R., Fischer, J., 2021. Optimal estimation MSG-SEVIRI clear-sky total column water vapour retrieval using the split window difference. *Atmosphere* 12 (10), 1256. <https://doi.org/10.3390/atmos12101256>.
- Keys, R., 1981. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* 29 (6), 1153–1160. <https://doi.org/10.1109/TASSP.1981.1163711>.
- Kotchenova, S.Y., Vermote, E.F., Matarrese, R., Klemm Jr., F.J., 2006. Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: path radiance. *Appl. Opt.* 45 (26), 6762–6774. <https://doi.org/10.1364/AO.45.006762>.
- Lagouarde, J.P., Bhattacharya, B.K., Crebassol, P., Gamet, P., Babu, S.S., Boulet, G., Ramakrishnan, R., 2018, July. The Indian-French Trishna mission: Earth observation in the thermal infrared with high spatio-temporal resolution. In: *Igarss 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 4078–4081. <https://doi.org/10.1109/IGARSS.2018.8518720>.
- Laraby, K.G., Schott, J.R., 2018. Uncertainty estimation method and Landsat 7 global validation for the Landsat surface temperature product. *Remote Sens. Environ.* 216, 472–481. <https://doi.org/10.1016/j.rse.2018.06.026>.
- Laraby, K.G., Schott, J.R., Raqueno, N., 2016, May. Developing a confidence metric for the Landsat land surface temperature product. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII*, 9840. SPIE, pp. 59–72. <https://doi.org/10.1117/12.222582>.
- Lewis, A., Lacey, J., Mecklenburg, S., Ross, J., Siqueira, A., Killough, B., Hosford, S., 2018, July. CEOS analysis ready data for land (CARD4L) overview. In: *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 7407–7410. <https://doi.org/10.1109/IGARSS.2018.8519255>.
- Livo, K.E., Kruse, F.A., Clark, R.N., Kokaly, R.F., Shanks III, W.C., 2007. Hydrothermally Altered Rock and Hot-Spring Deposits at Yellowstone National Park—Characterized using Airborne Visible and Infrared-Spectroscopy Data. <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1073&context=usgspubs>.
- Loveland, T.R., Dwyer, J.L., 2012. Landsat: building a strong future. *Remote Sens. Environ.* 122, 22–29. <https://doi.org/10.1016/j.rse.2011.09.022>.
- Loveland, T.R., Irons, J.R., 2016. Landsat 8: the plans, the reality, and the legacy. *Remote Sens. Environ.* 185, 1–6. <https://doi.org/10.1016/j.rse.2016.07.033>.
- Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An operational land surface temperature product for Landsat thermal data: methodology and validation. *IEEE Trans. Geosci. Remote Sens.* 56 (10), 5717–5735. <https://doi.org/10.1109/TGRS.2018.2824828>.
- Markham, B., Helder, D., 2012. Forty-year calibrated record of earth-reflected radiance from Landsat: a review. *Remote Sens. Environ.* 122, 30–40. <https://doi.org/10.1016/j.rse.2011.06.026>.
- Martí-Cardona, B., Prats, J., Niclòs, R., 2019. Enhancing the retrieval of stream surface temperature from Landsat data. *Remote Sens. Environ.* 224, 182–191. <https://doi.org/10.1016/j.rse.2019.02.007>.
- Martin, M.A., Ghent, D., Pires, A.C., Götsche, F.M., Cermak, J., Remedios, J.J., 2019. Comprehensive in situ validation of five satellite land surface temperature data sets over multiple stations and years. *Remote Sens.* 11 (5), 479. <https://doi.org/10.3390/rs11050479>.
- Meerdink, S.K., Hook, S.J., Roberts, D.A., Abbott, E.A., 2019. The ECOSTRESS spectral library version 1.0. *Remote Sens. Environ.* 230, 111196. <https://doi.org/10.1016/j.rse.2019.05.015>.
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American regional reanalysis. *Bull. Am. Meteorol. Soc.* 87 (3), 343–360. <https://doi.org/10.1175/BAMS-87-3-343>.
- Moik, J.G., 1980. *Digital Processing of Remotely Sensed Images*, 431. Scientific and Technical Information Branch, National Aeronautics and Space Administration.
- Montanaro, M., Levy, R., Markham, B., 2014a. On-orbit radiometric performance of the Landsat 8 thermal infrared sensor. *Remote Sens.* 6 (12), 11753–11769. <https://doi.org/10.3390/rs61211753>.
- Montanaro, M., Gerace, A., Lunsford, A., Reuter, D., 2014b. Stray light artifacts in imagery from the Landsat 8 thermal infrared sensor. *Remote Sens.* 6 (11), 10435–10456. <https://doi.org/10.3390/rs61110435>.
- Montanaro, M., McCorkel, J., Tveekrem, J., Stauder, J., Mentzell, E., Lunsford, A., Reuter, D., 2022. Landsat 9 thermal infrared sensor 2 (TIRS-2) stray light mitigation and assessment. *IEEE Trans. Geosci. Remote Sens.* 60, 1–8. <https://doi.org/10.1109/TGRS.2022.3177312>.
- Montanaro, M., McCorkel, J., Pearlman, A., Efremova, B., Wenny, B., Lunsford, A., Jhabvala, M., 2024, October. Landsat 9 thermal infrared sensor 2 (TIRS-2) radiometric calibration and potential future Landsat thermal band efforts. *AIP Conf. Proc.* 3230 (1), 100006. AIP Publishing LLC. <https://doi.org/10.1063/5.0234843>.
- Moussavi, M., Pope, A., Halberstadt, A.R.W., Trusel, L.D., Cioffi, L., Abdalati, W., 2020. Antarctic supraglacial lake detection using Landsat 8 and Sentinel-2 imagery: towards continental generation of lake volumes. *Remote Sens.* 12 (1), 134. <https://doi.org/10.3390/rs12010134>.
- Nádudvari, Á., Abramowicz, A., Maniscalco, R., Viccaro, M., 2020. The estimation of lava flow temperatures using landsat night-time images: case studies from eruptions of Mt. Etna and Stromboli (Sicily, Italy), Kilauea (Hawaii Island), and Eyjafjallajökull and Holuhraun (Iceland). *Remote Sens.* 12 (16), 2537. <https://doi.org/10.3390/rs12162537>.
- NASA, 2024. International Space Station Deorbit Analysis Summary. Available at: <https://www.nasa.gov/wp-content/uploads/2024/06/iss-deorbit-analysis-summary.pdf> (accessed 27 March 2026).
- NASA Earthdata, 2025. Transition from MODIS to VIIRS. Available at: <https://www.earthdata.nasa.gov/data/alerts-outages/transition-from-modis-viirs> (accessed 26 March 2026).
- NASA JPL, 2014. ASTER Global Emissivity Dataset, 100-Meter, HDF5 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/COMMUNITY/ASTER_GED/AG100.003 (accessed 31 December 2025).
- Niclòs, R., Pérez-Planells, L., Coll, C., Valiente, J.A., Valor, E., 2018. Evaluation of the S-NPP VIIRS land surface temperature product using ground data acquired by an autonomous system at a rice paddy. *ISPRS J. Photogramm. Remote Sens.* 135, 1–12. <https://doi.org/10.1016/j.isprsjprs.2017.10.017>.
- Niclòs, R., Perelló, M., Puchades, J., Coll, C., Valor, E., 2023. Evaluating Landsat-9 TIRS-2 calibrations and land surface temperature retrievals against ground measurements using multi-instrument spatial and temporal sampling along transects. *Int. J. Appl. Earth Obs. Geoinf.* 125, 103576. <https://doi.org/10.1016/j.jag.2023.103576>.
- NOAA National Data Buoy Center, 2026. Dataset. Available at: <https://www.ndbc.noaa.gov/> (accessed 20 April 2026).
- Padula, F.P., 2008. Historic Thermal Calibration of Landsat 5 TM Through an Improved Physics Based Approach. Rochester Institute of Technology. <https://repository.rit.edu/cgi/viewcontent.cgi?article=3965&context=theses> (accessed 20 May 2026).
- Pearlman, A., McCorkel, J., Montanaro, M., Efremova, B., Wenny, B., Lunsford, A., Reuter, D., 2018, September. Landsat 9 thermal infrared sensor 2 pre-launch characterization: initial imaging and spectral sampling along transects. In: *Earth Observing Systems XXIII*, 10764. SPIE, pp. 30–39. <https://doi.org/10.1117/12.2321084>.
- Qin, Z., Dall’Omo, G., Karnieli, A., Berliner, P., 2001. Derivation of split window algorithm and its sensitivity analysis for retrieving land surface temperature from NOAA-advanced very high resolution radiometer data. *J. Geophys. Res. Atmos.* 106 (D19), 22655–22670.
- Rabuffi, F., Hulley, G., Hook, S.J., Cawse-Nicholson, K., Ramsey, M.S., Thompson, J.O., La, T.T., 2025. Surface mineralogy using thermal infrared spectroscopy data from ECOSTRESS and ASTER. *IEEE Geosci. Remote Sens. Lett.* <https://doi.org/10.1109/LGRS.2025.3625069>.
- Radeloff, V.C., Roy, D.P., Wulder, M.A., Anderson, M., Cook, B., Crawford, C.J., Zhu, Z., 2024. Need and vision for global medium-resolution Landsat and Sentinel-2 data products. *Remote Sens. Environ.* 300, 113918. <https://doi.org/10.1016/j.rse.2023.113918>.
- Ramsey, M.S., Christensen, P.R., 1998. Mineral abundance determination: quantitative deconvolution of thermal emission spectra. *J. Geophys. Res.* 103 (B1), 577–596. <https://doi.org/10.1029/97JB02784>.
- Reek, J.E., Crowther, T.W., Lauber, T., Schemm, S., Parastatidis, D., Chrysoulakis, N., Smith, G.R., 2025. Forest edges are globally warmer than interiors and exceed optimal temperatures for vegetation productivity. *Commun. Earth Environ.* 6 (1), 635. <https://doi.org/10.1038/s43247-025-02626-1>.
- Reuter, D.C., Richardson, C.M., Pellerano, F.A., Irons, J.R., Allen, R.G., Anderson, M., Thome, K.J., 2015. The thermal infrared sensor (TIRS) on Landsat 8: design overview and pre-launch characterization. *Remote Sens.* 7 (1), 1135–1153. <https://doi.org/10.3390/rs70101135>.
- Rienecker, M.M., Suarez, M.J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.C., Nielsen, J.E., 2008. The GEOS-5 Data Assimilation System-Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0 (No. NASA/TM-2008-104606-VOL-27). Available online: <https://ntrs.nasa.gov/citations/20120011955>.
- Roy, D.P., Wulder, M.A., Gorelick, N., Hansen, M., Healey, S., Hostert, P., Zhu, Z., 2026. The next Landsat: mission turning point? *Remote Sens. Environ.* 332, 115087. <https://doi.org/10.1016/j.rse.2025.115087>.
- Rozenstein, O., Qin, Z., Derimian, Y., Karnieli, A., 2014. Derivation of land surface temperature for Landsat-8 TIRS using a split window algorithm. *Sensors* 14 (4), 5768–5780. <https://doi.org/10.3390/s140405768>.
- Saunders, R., Rayer, P., Brunel, P., Von Engeln, A., Bormann, N., Strow, L., Turner, D.S., 2007. A comparison of radiative transfer models for simulating atmospheric infrared sounder (AIRS) radiances. *J. Geophys. Res. Atmos.* 112 (D1). <https://doi.org/10.1029/2006JD007088>.
- Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, B., Brunel, P., Lupu, C., 2018. An update on the RTTOV fast radiative transfer model (currently at version 12). *Geosci. Model Dev.* 11 (7), 2717–2737. <https://doi.org/10.5194/gmd-11-2717-2018>.
- Scambos, T.A., Hulbe, C., Fahnestock, M., Bohlander, J., 2000. The link between climate warming and break-up of ice shelves in the Antarctic Peninsula. *J. Glaciol.* 46 (154), 516–530. <https://doi.org/10.3189/172756500781833043>.
- Scambos, T.A., Shuman, C., Fahnestock, M., Snow, T., Crawford, C.J., 2024. Monitoring polar ice change in the twilight zone. *Eos Am. Geophys. Union.* <https://doi.org/10.1029/2024EO240048>.
- Schaeffer, B.A., Iames, J., Dwyer, J., Urquhart, E., Salls, W., Rover, J., Seegers, B., 2018. An initial validation of Landsat 5 and 7 derived surface water temperature for US lakes, reservoirs, and estuaries. *Int. J. Remote Sens.* 39 (22), 7789–7805. <https://doi.org/10.1080/01431161.2018.1471545>.
- Schott, J.R., 1986. The role of remotely sensed data in studies of the thermal bar. *Remote Sens. Rev.* 1 (2), 341–358. <https://doi.org/10.1080/02757258609532073>.
- Schott, J.R., Hook, S.J., Barsi, J.A., Markham, B.L., Miller, J., Padula, F.P., Raqueno, N. G., 2012. Thermal infrared radiometric calibration of the entire Landsat 4, 5, and 7 archive (1982–2010). *Remote Sens. Environ.* 122, 41–49. <https://doi.org/10.1016/j.rse.2011.07.022>.
- Senay, G.B., Bohms, S., Singh, R.K., Gowda, P.H., Velpuri, N.M., Alemu, H., Verdin, J.P., 2013. Operational evapotranspiration mapping using remote sensing and weather datasets: a new parameterization for the SSEB approach. *J. Am. Water Resour. Assoc.* 49 (3), 577–591. <https://doi.org/10.1111/jawr.12057>.
- Senay, G.B., Parrish, G.E., Schauer, M., Friedrichs, M., Khand, K., Boiko, O., Ji, L., 2023. Improving the operational simplified surface energy balance evapotranspiration

- model using the forcing and normalizing operation. *Remote Sens.* 15 (1), 260. <https://doi.org/10.3390/rs15010260>.
- Shreevastava, A., Hulley, G., Prasanth, S., Chakraborty, T.C., Ramos Aguilera, D., Twomey Sanders, K., Yin, Y., 2025. Contemporary income inequality outweighs historic redlining in shaping intra-urban heat disparities in Los Angeles. *Nat. Commun.* 16 (1), 4950. <https://doi.org/10.1038/s41467-025-59912-x>.
- Siqueira, A., Lewis, A., Thankappan, M., Szantoi, Z., Goryl, P., Labahn, S., Lacey, J., 2019, July. CEOS analysis ready data for land—an overview on the current and future work. In: *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 5536–5537. <https://doi.org/10.1109/IGARSS.2019.8899846>.
- Southworth, C.S., 1983. General Characteristics and Availability of Landsat 3 and Heat Capacity Mapping Mission Thermal Infrared Data: U.S. Geological Survey Open-File Report, pp. 83–123. <https://doi.org/10.3133/ofr83123>.
- Thenkabail, P.S., Hanjra, M.A., Dheeravath, V., Gumma, M., 2010. A holistic view of global croplands and their water use for ensuring global food security in the 21st century through advanced remote sensing and non-remote sensing approaches. *Remote Sens.* 2 (1), 211–261. <https://doi.org/10.3390/rs2010211>.
- Theocharous, E., Fox, N.P., Barker-Snook, I., Niclòs, R., Santos, V.G., Minnett, P.J., Donlon, C.J., 2019. The 2016 CEOS infrared radiometer comparison: part II: laboratory comparison of radiation thermometers. *J. Atmos. Ocean. Technol.* 36 (6), 1079–1092. <https://doi.org/10.1175/JTECH-D-18-0032.1>.
- Thome, K., Reuter, D., Lunsford, A., Montanaro, M., Smith, R., Tesfaye, Z., Wenny, B., 2011, July. Calibration of the thermal infrared sensor on the Landsat data continuity mission. In: *2011 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 985–988. <https://doi.org/10.1109/IGARSS.2011.6049298>.
- Thompson, D.R., Kahn, B.H., Brodrick, P.G., Lebsock, M.D., Richardson, M., Green, R.O., 2021. Spectroscopic imaging of sub-kilometer spatial structure in lower-tropospheric water vapor. *Atmos. Meas. Tech.* 14 (4), 2827–2840.
- Tukey, J.W., 1977. *Exploratory Data Analysis*. Addison-Wesley.
- Turpie, K.R., Casey, K.A., Crawford, C.J., Guild, L.S., Kieffer, H., Lin, G., Thome, K., 2023. Calibration and validation for the surface biology and geology (sbg) mission concept: recommendations for a multi-sensor system for imaging spectroscopy and thermal imagery. *J. Geophys. Res. Biogeosci.* 128 (9), e2023JG007452. <https://doi.org/10.1029/2023JG007452>.
- USGS, 2025. Landsat 8-9 Calibration and Validation Algorithm Description Document, Version 5.0. <https://www.usgs.gov/media/files/landsat-8-9-calibration-validation-algorithm-description-document> (Accessed on 1 January 2026).
- USGS Landsat Missions, 2022. Landsat Collection 2 Known Issues. <https://www.usgs.gov/landsat-missions/landsat-collection-2-known-issues#ST> (accessed March 26, 2026).
- Vanhellemont, Q., Brewin, R.J., Bresnahan, P.J., Cyronak, T., 2022. Validation of Landsat 8 high resolution sea surface temperature using surfers. *Estuar. Coast. Shelf Sci.* 265, 107650. <https://doi.org/10.1016/j.ecss.2021.107650>.
- Vaughan, R.G., Hook, S.J., 2006. Using satellite data to characterize the temporal thermal behavior of an active volcano: Mount St. Helens, WA. *Geophys. Res. Lett.* 33 (20). <https://doi.org/10.1029/2006GL027957>.
- Vaughan, R.G., Hook, S.J., Calvin, W.M., Taranik, J.V., 2005. Surface mineral mapping at Steamboat Springs, Nevada, USA, with multi-wavelength thermal infrared images. *Remote Sens. Environ.* 99 (1–2), 140–158. <https://doi.org/10.1016/j.rse.2005.04.030>.
- Vlassova, L., Pérez-Cabello, F., Mimbreno, M.R., Llovería, R.M., García-Martín, A., 2014. Analysis of the relationship between land surface temperature and wildfire severity in a series of Landsat images. *Remote Sens.* 6 (7), 6136–6162. <https://doi.org/10.3390/rs6076136>.
- Wan, Z., 2008. New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sens. Environ.* 112 (1), 59–74. <https://doi.org/10.1016/j.rse.2006.06.026>.
- Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* 34 (4), 892–905. <https://doi.org/10.1109/36.508406>.
- Wan, Z., Li, Z.L., 2008. Radiance-based validation of the V5 MODIS land-surface temperature product. *Int. J. Remote Sens.* 29 (17–18), 5373–5395. <https://doi.org/10.1080/01431160802036565>.
- Wang, K., Liang, S., 2009. Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. *Remote Sens. Environ.* 113 (7), 1556–1565. <https://doi.org/10.1016/j.rse.2009.03.009>.
- Wenny, B.N., Helder, D., Hong, J., Leigh, L., Thome, K.J., Reuter, D., 2015. Pre-and post-launch spatial quality of the Landsat 8 thermal infrared sensor. *Remote Sens.* 7 (2), 1962–1980. <https://doi.org/10.3390/rs70201962>.
- Wilson, R.C., Hook, S.J., Schneider, P., Schladow, S.G., 2013. Skin and bulk temperature difference at Lake Tahoe: a case study on lake skin effect. *J. Geophys. Res. Atmos.* 118 (18), 10–332. <https://doi.org/10.1002/jgrd.50786>.
- Wu, Z., Snyder, G., Vadnais, C., Arora, R., Babcock, M., Stensaas, G., Newman, T., 2019. User needs for future Landsat missions. *Remote Sens. Environ.* 231, 111214. <https://doi.org/10.1016/j.rse.2019.111214>.
- Wulder, M.A., Roy, D.P., Radeloff, V.C., Loveland, T.R., Anderson, M.C., Johnson, D.M., Cook, B.D., 2022. Fifty years of Landsat science and impacts. *Remote Sens. Environ.* 280, 113195. <https://doi.org/10.1016/j.rse.2022.113195>.
- Xie, J., Lee, Z., Li, X., Wang, D., Zhang, C., Wu, Y., Yu, X., Zheng, Z., 2024. Estimation of sea surface temperature from Landsat-8 measurements via neural networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* <https://doi.org/10.1109/JSTARS.2024.3453908>.
- Yamaguchi, Y., Kahle, A.B., Tsu, H., Kawakami, T., Pniel, M., 1998. Overview of advanced spaceborne thermal emission and reflection radiometer (ASTER). *IEEE Trans. Geosci. Remote Sens.* 36 (4), 1062–1071. <https://doi.org/10.1109/36.700991>.
- Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., Sobrino, J.A., 2018. Satellite remote sensing of surface urban heat islands: progress, challenges, and perspectives. *Remote Sens.* 11 (1), 48. <https://doi.org/10.3390/rs11010048>.
- Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* 159, 269–277. <https://doi.org/10.1016/j.rse.2014.12.014>.